UNIVERSITE KASDI MERBAH OUARGLA
Faculté des lettres et des Langues
Département des Lettres et des Langues Étrangères
École Doctorale de Français
Antenne de l'Université de Ouargla



Mémoire présenté pour l'obtention

Du diplôme de Magistère

Filière : Français

Option : Sciences du langage

<u>Titre</u>

La problématique de l'étranger et/ou de l'étrangeté en polysémie dans le traitement automatique du langage naturel(TALN)

(Cas d'articles traduits sur « Google translate »)

Présenté par :
Melle Malika Mestour

Dirigé par :
Dr Salah Khennour

Membres de jury

Foudil DAHOU	Professeur, Université de Ouargla	Président
Abdelouahab DAKHIA	Professeur, Université de Biskra	Examinateur
Gaouaou MANAA	Professeur, Université de Batna	Examinateur
Salah KHENNOUR	Maître de conférences, Université de Ouargla	Rapporteur

Année universitaire : 2012-2013

Je dédie ce travail

A mes très chers parents, qui m'ont toujours été présent pour m'encourager et me soutenir pendant toutes mes années d'études, aucune autre personne ne m'est aussi trop chère que vous. Que dieu vous bénisse, vous assiste et vous soit en aide

A mes frères et sœurs que j'aime beaucoup et qui sont toujours là pour moi, Djamila, Karima, Mohamed, Hamid, Saliha, Souad et Lynda que dieu les garde pour moi

A tous mes neveux, Mélissa, Ines, Sarah, Amine et Fares

A mes meilleures amies Tina, Dihya, et Mira

A mon cher ami Mani pour toute l'aide qu'il m'a apporté

A Riad et toute la famille Medini

A tous mes amis de Ouargla : Mourad, Baya, Lamia, Salim, Fateh, Shahrazed, Nadra, Nadjla, Hichem, et mon cher frère Youcef

Remerciements

En premier lieu, nous remercions **Dieu,** le **Tout-Puissant** pour ses faveurs et ses grâces, de nous avoir donné le courage et la patience de mener ce travail durant toute cette année.

De plus, nos remerciements s'adressent à Mr **Salah Khennour**, pour nous avoir encadré et qui nous a apporté son expertise avec la plus grande disponibilité tout au long de l'année et ses conseils et pertinentes remarques.

Nos remerciements s'adressent aussi aux membres de jury pour avoir bien voulu accepter d'évaluer notre modeste travail et à tous nos enseignants de l'UKMO.

Nous tenons aussi à remercier toutes les personnes qui nous ont aidés pour bien mener ce travail. Nous remercions Mme **Aini Betouche** et Mme **Lila Abedslam** ainsi que tous
Les enseignants de l'UMMTO.

Nos remerciements vont également à Mr **Guy Lapalme**, et Mr **Daniel Valois** pour toute l'aide qu'ils nous ont accordée tout au long de notre séjour à Montréal.

Sommaire

Introduction générale………………………………………………………02

Partie théorique
Considérations théoriques

Chapitre I : Outil informatique et traitement automatique des langues des langues

Préambule 12
I. Présentation du traitement automatique du langage naturel TALN 12
I.1. Bref historique du traitement automatique du langage naturel 13
I.2. Disciplines regroupées en TALN 16
I.2.1 Linguistique 16
I.2.2 Intelligence artificielle 17
I.2.3 Informatique 17
I.2.4 Linguistique informatique 18
I.3 Les domaines du traitement automatique des langues 19
I.3.1 Traitement de la parole 19
I.3.2 Compréhension automatique des textes 19
I.3.3 Génération automatique des textes 20
I.3.4 Dialogue homme-machine 20
II. La traductique 20
II.1. La traduction 20
II.1.1 L’équivalence en traduction 21
II.1.2 Les différents types d’équivalence 22
II.1.2.1 L’équivalence linguistique 22
II.1.2.2 L’équivalence paradigmatique 22
II.1.2.3 Equivalence stylistique 22
II1.2.4 Equivalence sémantique 23
II1.2.5 Equivalence formelle 23
II.1.3 La notion d’interprétation 23

II-2-Premiers pas en traduction automatique 24
II-2-1-Les systèmes de traduction automatique 26
II-2-1-1-Les systèmes directs 26
II-2-1-2-Les systèmes de transfert 27
II-2-2-Mémoire de traduction 27

Chapitre II : Les théories linguistiques au service de la polysémie

I Le signe linguistique 30
I.1 Le signe linguistique selon Ferdinand De Saussure 30
I.2 D'autres définitions du signe linguistique 32
II Sens, signification, référent et référence 35
II.1 Définition du sens 35
II.2 Sens et signification 37
II.3 Sens, référent et référence 37
III La polysémie : une première définition 38
III.1 Définition de la polysémie selon la sémantique puissancielle 41
III.2 Polysémie vs homonymie 42
III.3 Les types de polysémie 44
III.3.1 Polysémie verticale 44
III.3.2 Polysémie logique 45
III.3.3 Polysémie contextuelle 45
III.4 Polysémie et traitement automatique des langues 46
III.5 Polysémie et traduction 47
Synthèse 48

PARTIE PRATIQUE
LE TALN : VERS UNE IMPOSTURE DES MOTS

Préambule 51

Présentation du corpus

I. Google translate **51**
II. L'article journalistique 53
III. Langue de traduction 54
I.1. La langue française 54
I.2. La langue arabe standard 55
IV. La traduction des articles sur Google translate 56
I. Article n 1 56
II. Article n 2 61

Chapitre I : L'existence du sens et/ou des sens

I Polysémie : passage de la signification au sens 66
I.1 Dénotation et connotation 66
I.2 Détermination du sens polysémique 67
I.2.1 Polysémie et contexte 67
I.2.2 Polysémie et cotexte 69
II Recueille des sens polysémique dans le corpus 71
II.1 Article I 71
II.2 Article II 73
III Le sens d'un énoncé 75
IV Étranger et étrangeté selon Jean Claude Cauquet 78

Chapitre II : D'une langue étrangère à l'étrangeté du langage

I La désambiguïsation sémantique 80
II Google traduction : comment ça marche ? 81
II.1 Calcul de sens polysémique sur Google traduction 83
II.2 La méthode statistique 83
II.2.1 Corpus parallèles 84

II.2.2 L'opération de traduction..85

II.2.3 Modèle probabiliste du langage 87

III Problèmes de traduction sur Google 89

III.1 Problèmes liés au contexte dans le processus traductique 90

III.2 Problèmes liés au cotexte dans le processus traductique 90

III.3 D'une langue étrangère à l'étrangeté du langage...92

Synthèse………………………………………………………………………………92

Conclusion générale ………………………………………………………..95

Bibliographie……………………………………………………………103

Résumé

Le traitement automatique du langage naturel (TALN) est un domaine nouveau en linguistique. L'élaboration de systèmes informatiques intelligents capables de reproduire le langage humain est l'objectif assigné au TALN. Ainsi, la traduction automatique est la première branche du TALN où nous avons essayé d'automatiser cette faculté propre à l'Homme. En revanche, ce domaine reste difficile à cause des ambiguïtés inhérentes aux phénomènes linguistiques, telle la polysémie qui reste la plus handicapante. Elle est un fait important dans toutes les langues ; elle présente pour un même signe linguistique plusieurs sens. Ainsi, Nous allons, dans notre étude, essayer de démontrer si le TALN, dans les articles de presses traduits sur « Google translate », mène à l'étranger et/ou l'étrangeté du langage et ceci par le biais des unités polysémiques.

Mots clés: sens, polysémie, étranger, étrangeté, TALN.

Summary

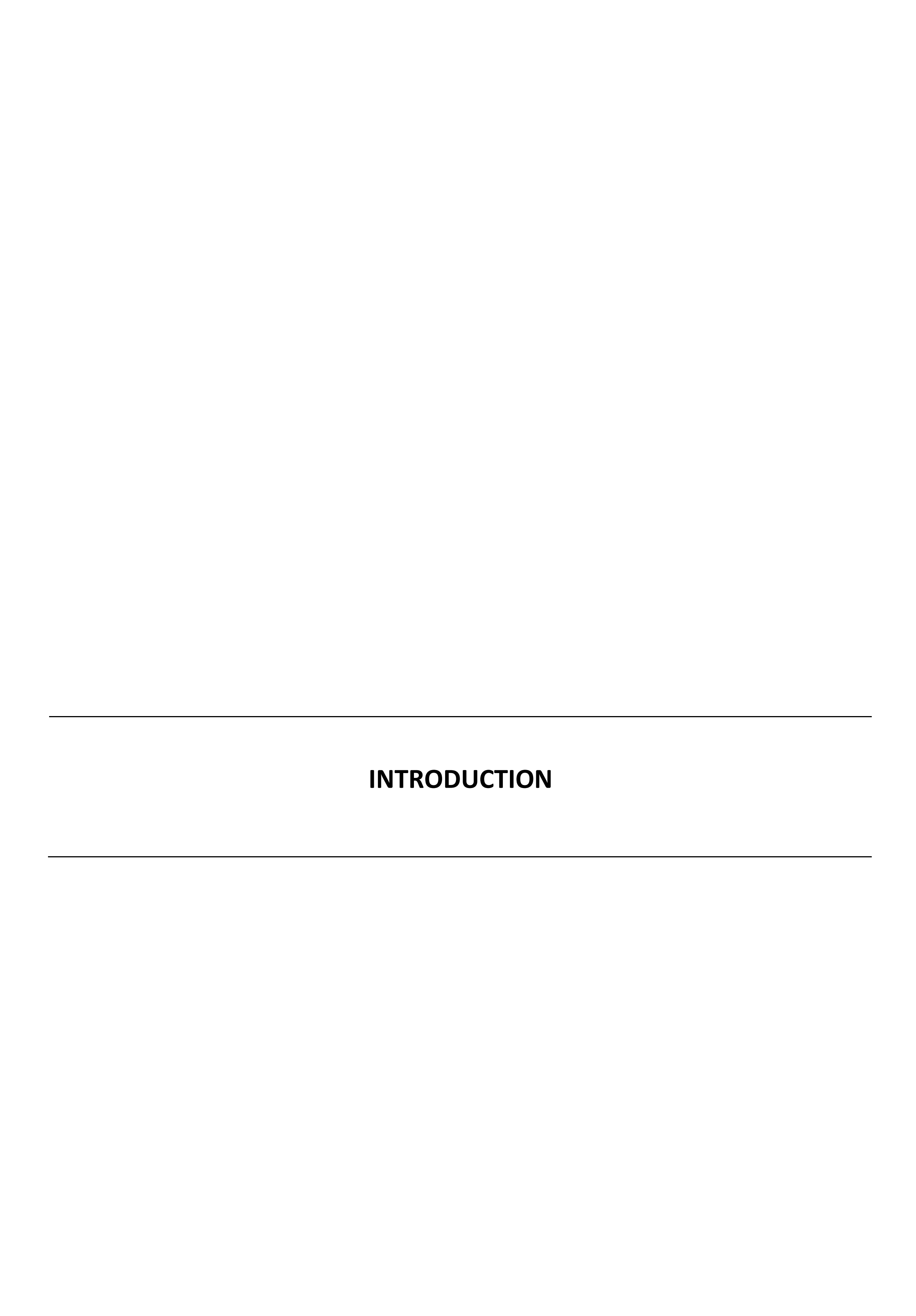
The natural language processing (NLP) is a new field in linguistics. The development of intelligent computer systems that can replicate human speech is the objective assigned to NLP. Thus, machine translation is the first branch of NLP where we tried to automate the faculty's own rights. However, this area remains difficult because of ambiguities inherent linguistic phenomena, such as polysemy remains the most debilitating. It is an important fact in all languages; it has the same sign for several linguistic meaning. Thus, we will, in our study, try to demonstrate whether the NLP in the press articles translated to "Google translate" leads abroad and / or the strangeness of the language and through these polysemic units.

Keywords: meaning, polysemy, abroad, strangeness, NLP.

ملخص

هو حقل جديد في علم اللغة. تطوير أنظمة الكمبيوتر الذكية التي يمكن تكرار كلمة الإنسان هو الهدف المخصصة وهكذا، الترجمة الآلية هي فرع الأول من حيث البرمجة اللغوية العصبية حاولنا أتمتة كلية الحقوق الخاصة. ومع .NLP ل ذلك، لا تزال هذه المنطقة صعبة بسبب الغموض اللغوي الأصيل الظواهر، مثل تعدد المعاني لا يزال معظم المنهكة. ومن الحقائق الهامة في جميع اللغات، فقد علامة نفس المعنى اللغوي لعدة. وبالتالي، فإننا، في دراستنا، في محاولة لإثبات ما إذا في المقالات الصحفية ترجم إلى "ترجمة جوجل" في الخارج يؤدي و / أو غرابة اللغة وخلال هذه الوحدات NLP كان polysemic.

كلمات البحث: معنى، تعدد المعاني في الخارج، الغرابة



INTRODUCTION

Le phénomène de la mondialisation des échanges a suscité le besoin de communication interlinguistique et interculturelle entre individus, entre sociétés et différents groupes. Au XXème siècle, les nouvelles technologies de l'information et de la communication (NTIC) ont connu un essor sans précédent. Des réseaux de communication tel Internet donnent accès à des sources d'informations autrefois inaccessibles. Cette révolution dans la production et la diffusion de l'information a surmonté les barrières de distance et des frontières à la fois géographiques et linguistiques. Cela a favorisé la convergence des disciplines.

Le traitement automatique du langage naturel (TALN) se trouve au cœur des débats. Ce nouveau domaine est le fruit d'un croisement entre l'informatique (relevant des sciences dures) et la linguistique (des sciences humaines). Ces sciences semblaient auparavant séparables et loin d'avoir le même champ d'intérêt. Dans les pays développés, d'ailleurs, des disciplines hybrides issues de celles déjà connues s'enseignent dans les universités. Pour cette raison, ce travail à entreprendre sera une contribution à la création de disciplines issues d'un croisement entre l'informatique avec ce qu'elle comporte comme fenêtres ouvertes à la Toile et la linguistique dans sa dynamique significative (la sémantique/sémiotique).

Le TALN en est un exemple relevant du domaine de recherche pluridisciplinaire. Il regroupe, outre l'informatique et la linguistique, les mathématiques, les sciences cognitives, les statistiques et l'intelligence artificielle. Ainsi, l'élaboration de systèmes informatiques intelligents capables de reconnaitre, de comprendre, d'interpréter, et de reproduire le langage humain est l'objectif assigné au TALN.

Historiquement, la traduction automatique est le premier domaine du TALN. C'est au cours de la guerre froide que nous assistons à la première traduction automatique. Quelques phrases russes, sélectionnées à l'avance, furent traduites en anglais. Cette expérience a déclenché de nombreuses recherches dans le domaine.

Par son qualificatif « naturel », le langage humain est difficile à la manipulation algorithmique. De nombreuses ambiguïtés, dues aux différents sens attribués aux mots à chaque fois que nous changeons de cotexte et de contexte, expliquent la complexité de ce langage dit « naturel ».

Ainsi, le sens linguistique a toujours fait l'objet d'étude des linguistes, et aujourd'hui, il constitue le problème fondamental en TALN. Pour Saussure, le sens est *l'entité du signifié*,

c'est *« l'image conceptuelle du signe linguistique ».*[1] Les linguistes A.J Greimas et J. Courtés affirment cependant que le sens est *« indéfinissable »*[2].

Parmi les problèmes liés au sens, nous avons choisi de travailler sur la polysémie. Elle est un fait linguistique, un trait fort répondu et important de toutes les langues naturelles. Le préfixe « poly », d'origine grecque, qui signifie « nombreux », et le radicale « sème » qui signifie « sens » nous révèlent une première définition de la polysémie. C'est la « *propriété d'un mot qui présente plusieurs sens.* »[3], ou « *la propriété d'un signe linguistique qui a plusieurs sens.* ».[4]

En outre, G. Kleiber la définit comme une « *pluralité de sens liée a une seule forme, des sens qui ne paraissent pas totalement disjoints, mais qui se trouvent unis par tel ou tel rapport. »*[5]. Par ce passage, un autre critère vient s'ajouter à la définition déjà donnée. La polysémie présente pour un même *signifiant* plusieurs *signifiés* liés entre eux. Dans le langage courant, nous parlons de sens propre et de sens figuré. Des linguistes comme Jean Michel Adam parlent de sens dénoté et de sens connoté(s). En sémantique, nous considérons qu'il y a une stabilité sémantique et une dynamique sémique. Lors du processus de traduction via « Google translate », quel sens pourrions-nous par conséquent obtenir ? Un sens propre ou un sens figuré ? De toute évidence, la machine propose un sens. Ce sens correspond parfois à celui recherché. D'autres fois, le mot obtenu exprime une réalité tout autre au point d'émettre l'hypothèse de l'étrangeté de la langue.

Or, « Etrangeté »[6] et « Etranger » sont deux concepts de Jean-Claude Coquet, auteur de la théorie des instances énonçantes. Dans son acte, au colloque *Théorie littéraire et traduction* qui a eu lieu à l'INALCO, Coquet parle de l'étrangeté dans le langage. Pour lui *« le problème de la traduction est moins de savoir rendre compte de l'étranger, de ce qui dans le texte, écrit dans une langue donnée, et l'indice d'une autre culture que d'accéder à l'étrangeté. »*[7], c'est-à-dire le rôle du traducteur n'est pas de traduire la langue de l'autre vers une autre langue, mais sa tâche est de construire un sens, un contenu conceptuel. La traduction ne doit pas porter sur les langues mais sur le contenu du discours ou des textes.

---

[1] Ferdinand, De Saussure. *Cours de linguistique générale.* , Lausanne, Payot, 1995, p.108.
[2] A.J, Greimas et J, Courtés. *Sémiotique Dictionnaire raisonné de la théorie du langage*. Paris, Hachette, 1979, p.348.
[3]- *Le Petit Larousse illustré,* Larousse, Paris, 1991, p.778.
[4]-J. DUBOIS. Dictionnaire de linguistique et des sciences du langage, Larousse, Paris. 1994, pp.369-770.
[5]- G. KLEIBER, *Problèmes de sémantique : la polysémie en questions,* Presses Universitaires du Septen-Trion, Paris, 1999, p.157.
[6]http://books.google.fr/books?id=1umrXw4cz78C&pg=PA49&dq=s%C3%A9miotique+subjectale+coquet+traduction&hl=fr&ei=azrlTYqhM8ju0gGP4yfBw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC8Q6AEwAA#v=onepage&q=s%C3%A9miotique%20subjectale%20coquet%20traduction&f=false
[7]- Idem.

Or, cela est-il possible sans aboutir à « l'étrangeté » d'un langage qui échapperait à la compréhension ?

Dans le monde du partage multilingue de l'information, la traduction automatique suscite des besoins dans le monde. Traduire, c'est ce passage, par le biais des signes, d'une langue à une autre, et où chaque langue est un univers qui a une vision du monde et une culture propre. Or, le sens est l'unité du langage la plus difficile d'accès pour une étude scientifique et empirique. Il n'est ni observable, ni mesurable. Par conséquent, qualifier le nom « traduction » d'« automatique » pose de nombreux problèmes conceptuels pour avancer dans ce domaine.

Notre intérêt est porté sur la problématique de l'étranger et/ou l'étrangeté dans la polysémie en TALN, parce que, d'une part, la polysémie est l'un des aspects du langage qui nous affecte le plus dans notre vie quotidienne. N'est-ce pas que nous cherchons toujours à comprendre ce que l'autre voulait dire par tel ou tel mot ? En outre, par sa complexité, la polysémie a constitué et constitue toujours l'objet d'étude des linguistes. En effet, nombreux sont les ouvrages linguistiques traitant les problèmes de la polysémie. Picoche atteste que nous ne pouvons pas avoir une langue sans polysémie et que dans celle-ci, il y a un mode d'emploi qu'il faut découvrir : « *le langage ne pourrait pas se passer de la polysémie, elle est un caractère fondamental du lexique, les polysèmes sont de puissantes machines sémantiques extrêmement perfectionnées, servant à appréhender de vaste secteur du réel, et…l'on passe à coté des vrais problèmes si on ne cherche pas à comprendre comment elle fonctionne. »*[8].

D'autre part, le traitement automatique du langage naturel est un domaine récent en linguistique. Vouloir modéliser le langage humain à des fins d'automatisation nécessite beaucoup d'intérêt. C'est un domaine qui reste difficile à cause des ambiguïtés inhérentes aux phénomènes linguistiques. Ces ambiguïtés sont diverses, mais la plus handicapante est la polysémie.

Les difficultés majeures rencontrées par les systèmes de traduction naissent des problèmes de connotations qui mettent en cause les transferts de culture à culture, et de vision du monde à vision du monde, surtout lorsque les deux langues mises face à face n'ont aucun lien tels l'arabe et le français.

---

[8] - J. Picoche, *Structures sémantiques du lexique français*, Editions Nathan, Paris, 1986, P.3.

Les textes scientifique et technique sont monosémiques. Ils renvoient à une réalité objectivable. C'est une source de savoir vérifiable, c'est-à-dire qu'ils ont un contenu essentiellement informatif, dénotatif, donnant lieu à une seule interprétation possible. Ce type de texte ne pose pas de problèmes d'automatisation, et le système canadien Traduction Automatique de l'Université de Montrèal- Météo (TAUM-Météo)* illustre nos propos. Contrairement aux textes scientifique et technique, un texte littéraire dépeint une expérience en grande partie subjective, c'est un texte connotatif où nous constatons une forte présence de mots polysémiques.

Pour notre étude, le corpus à partir duquel nous allons travailler est un texte journalistique. Le message écrit de la presse est plus proche de la parole vécue immédiate. Etant lectrice de journaux et utilisatrice d'Internet, nous avons essayé à plusieurs reprises de traduire des textes journalistiques. Nous avons remarqué une forte présence de connotation qui pose problème au système « *Google translate* »[9]. Cette connotation s'explique par l'espace étroit qui pousse le journaliste à effectuer un choix dans son lexique afin de dire plusieurs choses avec un nombre restreint de mots.

Il existe sur le web plusieurs systèmes de traduction, mais nous avons choisi de mener notre étude sur le système « *Google translate* » parce qu'il est le système le plus avancé en matière de traduction. Contrairement aux autres systèmes, *Google* fonctionne par analogie statistique et l'entreprise compare d'importants corpus de textes de différentes langues, en l'occurrence l'arabe et le français. Par ailleurs, si notre recherche a visiblement une allure d'un travail à entreprendre en traduction, dans le fond, il ne sera pas le cas car nous nous attarderons plus à étudier les sèmes, les sémèmes à même de contribuer à l'élucidation du sens du texte à traduire.

Aussi, nous tenons à signaler que si nous avons choisi la traduction automatique, ce n'est pas parce que nous voulons nous inscrire dans le domaine de la traduction, mais après avoir fait plusieurs lectures et recherches dans le domaine du TALN, la traduction automatique est le moyen que nous avons jugé comme le plus utile pour bien mener notre étude. M. Wandruzka, dans son ouvrage *Vers une linguistique de la traduction* justifie notre choix : « *pour mieux comprendre la nature de nos langues humaines, c'est, en effet, dans la*

---

[9] http://translate.google.com/

*- TAUM-Météo est un système de traduction automatique conçu pour la traduction des bulletins météorologiques au Canada.

traduction qu'apparaissent au grand jour les convergences et les divergences entre les formes et les structures de nos langues particulières, que se dévoilent les universaux du langage humain. ».[10]

L'objectif poursuivi dans ce travail est de démontrer le rôle du contexte et du cotexte dans la compréhension d'une unité polysémique en TALN. Nous allons essayer d'expliquer comment, dans le cas de la traduction automatique, une automatisation du langage mène à une traduction de la langue et non pas du langage. Nous précisons que nous verrons les deux concepts « langue » et « langage » selon Jean-Claude Coquet. Ce dernier définit le langage comme un discours, c'est l'étrangeté, et la langue comme le produit du discours, c'est l'étranger[11].

Il est important de signaler que notre objectif est de soulever la problématique de l'étrangeté et/ou l'étranger soulevé par la polysémie en TALN. Nous n'allons pas essayer, dans notre étude, d'apporter du nouveau par rapport à ce qui a été déjà donné par les chercheurs en TALN. En effet, pour comprendre le sens d'une unité polysémique dans n'importe quel énoncé, nous avons recours au contexte pour enlever toute ambiguïté de sens. Pourtant, dès que l'on veut automatiser une telle performance, la polysémie devient un véritable problème et elle donne bien du souci aux chercheurs en TALN.

En guise de problématique, nous posons une série de questions auxquelles nous essayerons, par la suite, de proposer des hypothèses :

- Pouvons-nous regrouper toutes les connotations d'un polysème ? donc, pouvons-nous associer automatiquement un sens à un mot dans un énoncé ?
- Un ordinateur peut-il prendre en considération le contexte où apparaît le mot polysémique ?

- Une traduction automatique est-elle une traduction d'un sens global ou une traduction mot-à-mot ?

Répondre à toutes ces questions nous aidera à trouver une réponse à une question centrale : le traitement automatique de la polysémie a-t-il comme corollaire l'étranger et/ou

---

[10] M. Wandruzka, cité par. Yamina Hellal, *La théorie de la traduction: approche thématique et pluridisciplinaire*, office des Publications Universitaires, Alger, 2001, p.5.

[11] http://books.google.fr/books?id=1umrXw4cz78C&pg=PA49&dq=s%C3%A9miotique+subjectale+coquet+traduction&hl=fr&ei=azrlTYqhM8ju0gGP4yfBw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC8Q6AEwAA#v=onepage&q=s%C3%A9miotique%20subjectale%20coquet%20traduction&f=false

l'étrangeté dans le langage ? De ce fait, répondre à cette question nous permet d'atteindre l'objectif tracé dans cette étude.

Baylon et Fabre définissent la polysémie comme *« un phénomène diachronique qui consiste dans l'addition d'acceptions nouvelles au sens fondamental d'un signe. » et que « cette multiplication des sens aboutit, sur le plan synchronique, à la coexistence de plusieurs significations pour un même signe »*[12].

A partir de cette citation, nous pouvons postuler quelques hypothèses qui répondent à notre problématique. Nous pouvons recueillir, sur le plan synchronique, un nombre important d'acception du mot polysémique dans le but d'automatisation. Or, les langues évoluent et les polysèmes acquièrent de nouvelles acceptions une fois mise en contexte. Par conséquent, nous ne pouvons pas regrouper tous les sens connotés d'une unité polysémique.

Le décodage et l'interprétation de mots polysémiques demande au traducteur des connaissances qui sont extérieures au système même de la langue. Ainsi, pour réaliser un ordinateur tel HAL* a conduit les chercheurs à formaliser et enregistrer dans l'ordinateur toutes les connaissances linguistiques et extra-linguistiques. Cependant, la prise en compte par l'ordinateur des connaissances qui lui sont extérieures est difficile, voir impossible, parce qu'en premier lieu, ces connaissances sont vastes et changent, et en second lieu, un ordinateur est dépourvu de toute réflexion.

Par conséquent, vouloir automatiser les mots polysémiques mène à une traduction dépourvue de contexte, autrement dit, il sera question d'une traduction machinale, c'est-à-dire traduction mot-à-mot.

Ces hypothèses nous conduisent à conclure que le traitement automatique de la polysémie a comme corollaire une traduction de la langue, donc de l'étranger, et non pas de l'étrangeté dans le langage.

En vue de vérifier ces hypothèses et, de la sorte, répondre à notre problématique, nous avons établi une méthodologie s'articulant sur deux parties : une partie théorique portant sur des notions et définitions des concepts, et une partie pratique où nous analyserons notre corpus afin de répondre à nos interrogations.

---

[12]- CHR, BAYLON. & P. FABRE, *La sémantique*, Nathan, Paris, 1978, p.161.
*- Dans le célèbre film *« 2001, L'odyssée de l'espace »* de S. Kubrik, l'ordinateur HAL comprend l'homme, dialogue avec lui dans sa langue, exécute ses commandes et ressent des émotions.

Dans la partie théorique, nous allons présenter le premier chapitre s'intitulant : Outil informatique et traitement automatique des langues. Dans ce chapitre nous tenterons de faire un petit aperçu du croisement entre informatique et linguistique. Ensuite, nous expliquerons ce que la linguistique informatique et le traitement automatique du langage naturel. Enfin, nous expliquerons ce que la *traductique.*

Ensuite, nous passons à un deuxième chapitre s'intitulant : Les théories linguistiques au service de la polysémie. Une présentation théorique du signe linguistique nous permettra de mieux comprendre le concept de « polysémie ». Nous toucherons dans nos interrogations, entre autres aux notions suivantes : le sens, la signification, le référent et la référence. Nous expliquerons ce que le signifié de puissance et les signifiés d'effet selon Gustave Guillaume[13], et enfin nous verrons la place de la polysémie en traitement automatique du langage naturel ainsi qu'en traduction automatique.

Après une présentation des deux chapitres constituant la partie théorique, nous passerons à une partie pratique qui se divise, à son tour, en deux chapitres. Dans le premier chapitre, nous essayerons de démontrer, à partir de notre corpus, l'existence du sens et/ou des sens. Ce chapitre nous permettra de vérifier comment la traduction automatique mène d'une langue étrangère à l'étrangeté du langage. Ceci sera l'intitulé du second chapitre.

Une traduction d'un texte journalistique d'une langue à l'autre nécessite une approche contrastive. La comparaison des deux langues nous permettra de mieux comprendre le concept de « polysémie ». En revanche, la définition de la polysémie comme une pluralité de sens liés à un sens commun nous mènera à suivre l'approche connotative. Pour se faire, nous allons traduire un texte journalistique en nous basons sur le système « *Google translate* » de l'arabe vers le français. Cette traduction nous permettra de relever les différents mots polysémiques et, par la suite, nous expliquerons, à l'aide de l'approche connotative de Jean Michel Adam comment une traduction automatique mène à une traduction mot-à-mot, c'est un traitement du langage dépourvu du contexte, donc il mène à une étrangeté du langage selon le concept de Jean-Claude Coquet.

Dans son ouvrage *linguistique et discours littéraire,* Jean Michel Adam définit la connotation comme *« le produit d'une structuration textuelle du langage naturel »* et que *« le langage fonctionne dans un étroit rapport avec le domaine extralinguistique et, surtout,*

---

[13] G. GUILLAUME, *Langage et science du langage,* Librairie Nizet, Paris, 1994.

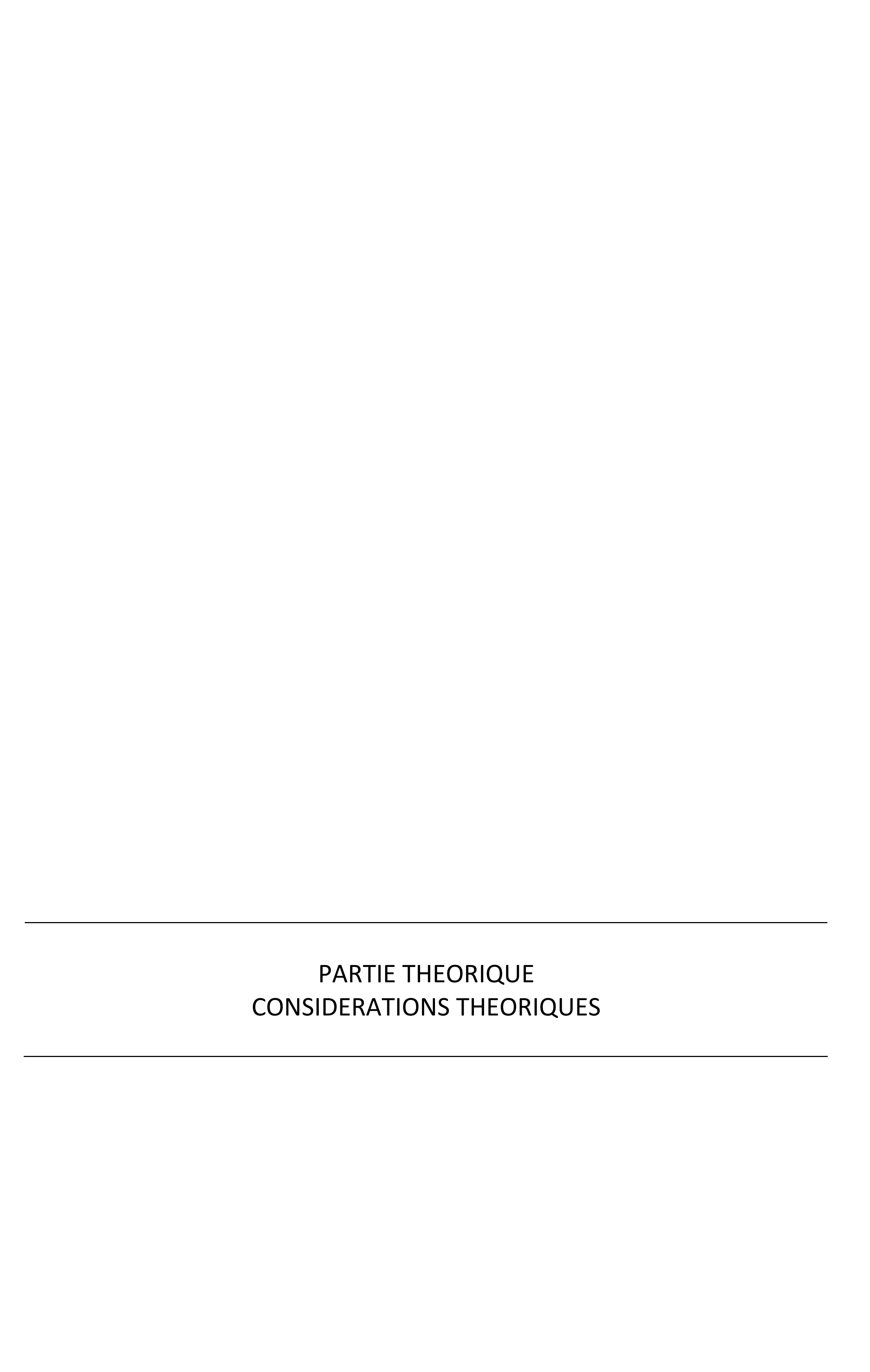
que les connotations d'un signe n'existent pas en elle-même : un dictionnaire des connotations est donc impensable »[14].

L'étude de la polysémie peut se faire sur deux plans : d'abord, nous pouvons l'aborder dans le domaine de la diachronie, c'est-à-dire regrouper tous les sens que peut avoir un mot polysémique depuis son étymon. Ensuite, nous pouvons étudier la polysémie dans le domaine de la synchronie, c'est-à-dire étudier les différents sens du mot polysémique à un moment donné.

Notre étude se place dans le domaine de la synchronie. La polysémie comme un phénomène diachronique ne constituera pas l'objet de notre intérêt, c'est pourquoi nous n'entamerons aucune étude diachronique des emplois des mots polysémiques. Nous souhaitons étudier que le français du XX siècle.

Nous espérons qu'une fois accomplie, cette recherche arguera en faveur d'un intérêt certain : celui de renseigner sur le traitement automatique du langage naturel, ainsi que les difficultés liés à ce domaine nouveau en linguistique. Nous espérons que notre étude sera par la suite une aide pour ceux qui veulent travailler dans le domaine de la linguistique informatique ou travailler sur la polysémie

[14] Jean Michel Adam, *Linguistique et discours littéraire : théorie et pratique des textes*, Larousse université, Canada, 1976, p.84.



PARTIE THEORIQUE
CONSIDERATIONS THEORIQUES

Préambule

Cette partie se donne pour objectif l'établissement du cadre théorique général à ce travail. L'étude des sens est compliquée par un fait essentiel du langage humain. Il n'existe pas une relation univoque et simple entre les formes et le sens. Plusieurs formes distinctes peuvent porter le même sens, mais une seule forme peut aussi porter plus d'un sens comme le cas de la polysémie.

Ainsi, en traitement automatique du langage naturel, en l'occurrence en traduction automatique, la polysémie est considérée comme un problème difficile à résoudre mais qui reste réalisable.

Nous allons essayer, dans cette partie, de définir quelques notions de bases. Nous essayerons, dans un premier chapitre, intitulé : « Outil informatique et traitement automatique des langues », de présenter le domaine du traitement automatique du langage naturel (TALN). Nous verrons les différents domaines de celle-ci ; et enfin nous exposerons ce qu'est la traductique avec ses principaux traits.

Par ailleurs, nous entamerons dans un second chapitre, intitulé « Les théories linguistiques au service de la polysémie », la notion de polysémie. Nous allons essayer de définir, en premier lieu, les notions de sens, signification, référent et référence. Ensuite, nous aborderons la notion de polysémie selon différents auteurs. Enfin, nous exposerons la notion de polysémie dans le domaine du traitement automatique du langage naturel.

Dans ce premier chapitre de la partie théorique, nous essayerons de présenter les grands traits du domaine de la linguistique informatique. Cette nouvelle discipline, connue aussi sous l'appellation de linguistique computationnelle, est issue des développements de l'informatique dans les domaines des sciences du langage.

Chapitre I

Outil informatique et traitement automatique des langues

I-Présentation du traitement automatique du langage naturel (TALN)

Dès la mise au point des ordinateurs, l'être humain a eu l'idée d'utiliser ces machines dans le but de traiter le langage naturel et aussi de traduire des textes en diverses langues. Mais qu'est ce que nous entendons par les deux termes « traiter » et « langage naturel » ?

En informatique, nous entendons par « traitement », la transformation d'un objet d'entré en un objet de sortie. Et en linguistique, nous appelons langage naturel la langue parlée et écrite par l'être humain par opposition aux langues artificielles comme les mathématiques, la logique et l'informatique.

Nous assistons, dès la seconde guerre mondiale, à l'essor d'un domaine nouveau, un domaine qui regroupe plusieurs disciplines telles que l'intelligence artificielle, l'informatique et la linguistique: nous parlons ici du Traitement Automatique du Langage Naturel. Or, qu'est ce que nous entendons par TALN ?

Nous regroupons sous le sigle TALN, l'ensemble des recherches et développement visant à modéliser et reproduire, à l'aide des ordinateurs, le langage naturel. Mais, pourquoi les chercheurs s'intéressent-ils autant à automatiser cette faculté propre à l'Homme ?

Vouloir automatiser le langage humain peut être justifié par deux sources de motivations. D'une part, le langage humain a toujours fasciné les chercheurs, en l'occurrence les chercheurs en intelligence artificielle (IA). Autrement dit, les chercheurs ont voulu comprendre le fonctionnement de ce langage dit naturel. D'autre part, disposer d'un ensemble d'outils automatiques facilitant la communication entre les quatre coins du monde est le but qui a poussé les chercheurs en TALN à vouloir automatiser le langage naturel.

Le traitement automatique du langage naturel TALN, s'intéresse aux traitements informatisés mettant en jeu du matériau linguistique. Parmi les traitements effectués dans ce domaine nous citons : l'analyse des textes, la génération de textes, correction orthographique et grammaticale et la traduction automatique.

I-1-Bref historique du traitement automatique du langage naturel

Historiquement, le traitement automatique du langage naturel remonte à la deuxième guerre mondiale. En 1954, nous assistons à la première traduction automatique. La guerre entre les américains et les russes a poussé les américains à inventer le premier traducteur automatique. Mais ce dernier était très rudimentaire. Ainsi, nous avons essayé de traduire quelques phrases russes vers l'anglais. Le nombre de mots n'était qu'à 250 et les règles grammaticales utilisées étaient au nombre de six.

Par la suite, Y. Bar-Hellal a organisé, en 1962, la première conférence en traduction automatique. Depuis cette époque, nous assistons à de nombreux travaux qui s'effectuent dans le domaine du TALN, et de lourds financements ont été investis.

Ainsi, la traduction d'une langue à une autre a été une traduction mot-à-mot. Pour cela, fabriquer et manipuler des dictionnaires électroniques étaient les principaux travaux tracés en TALN à cette période. Or, cette conception simplifiée de la traduction a donné beaucoup de phrases incorrectes, et les chercheurs se sont rendus compte que de nombreuses connaissances contextuelles, portant sur la situation décrite, et encyclopédiques, portant sur le monde, sont nécessaires pour trouver la bonne traduction d'un mot.

Dès lors, Le problème de la représentation des connaissances a été posé et nous avons eu le doute de la possibilité d'une traduction dite automatique. Ainsi, un groupe d'expert (l'Automatic Language Processing Advisory Council ALPAC) rédige un rapport où ils mentionnent que la traduction automatique coûte deux fois plus chère que la traduction humaine et donne des résultats moins bons. Tous ces résultats mènent à un échec de toute tentative de traduction automatique.

En revanche, malgré ces échecs, les années 50 voient l'apparition d'importantes publications. Ainsi, entre 1951 et 1954 Zellig Harris publie ses travaux les plus importants, *linguistique distributionaliste.* Ensuite, en 1957 N. Chomsky publie aussi d'importants travaux portant sur la syntaxe des langues naturelles, et sur les relations entre les grammaires formelles et les grammaires naturelles.

En outre, nous pouvons citer aussi, en 1956 à l'école de Dartmouth, la naissance de l'intelligence artificielle. Cette dernière part du point de vue que toute intelligence humaine peut être décrite de façon précise pour qu'une machine, comme l'ordinateur, le simule. Ainsi, les chercheurs les plus importants en intelligence artificielle à cette époque, tels que John Mc Carthy, Marvi Minsky, Allan NewelL, Herbert Simon ; parlent d'une possibilité de créer des programmes d'ordinateurs qui se comportent intelligemment, et en particulier qui soient capables d'utiliser le langage.

Des travaux encourageants ont été effectués par les chercheurs en intelligence artificielle et même par leurs élèves. Nous citons par exemple les élèves de Marvin Minsky, qui ont mis en œuvre des mécanismes de traitement simples mais importants. Leurs

résultats, en particulier le comportement spectaculaire d'ELIZA[15] qui simule un dialogue entre un psychiatre et son patient, relancent les recherches sur la compréhension automatique du langage. Mais ces systèmes ne fonctionnent que dans des situations de contexte restreint.

Nous citons également les travaux importants de Ross Quillian sur la représentation des connaissances. Il préconise l'utilisation de réseaux sémantiques afin de représenter le sens des mots et des phrases en explicitant les relations des divers concepts entre eux grâce à des liens qui indiquent le sens des relations. En plus, nous pouvons citer aussi Terry Winograd qui réalise en 1972, le premier logiciel SHRDLU qui est capable de dialoguer en anglais avec un robot.

Les années 70 voient le développement d'approches sémantiques. L'importance du contexte et le rôle essentiel d'une bonne connaissance du domaine traité pour comprendre un texte est ainsi mis en avant. De ce fait, les recherches ont cessé de se limiter à l'interprétation de phrases seules pour aborder le traitement d'unités plus importantes comme le récit et le dialogue. Ainsi, nous ne nous limitons plus au seul sens objectif et nous remarquons que la signification subjective dépend d'informations implicites qui font partie des connaissances générales communes aux interlocuteurs. M Minsky tente alors d'élaborer un cadre général de représentation des connaissances, *les formes*, alors que R. Schank s'efforce d'identifier clairement les diverses connaissances nécessaires dans un système interprétant le langage naturel.

D'un autre côté, des algorithmes sont proposés pour analyser les grammaires les plus simples et les modèles syntaxiques connaissent, de plus en plus, des développements en informatique. Or, ces développements étaient sous estimés parce que ces grammaires formalisées jusqu'ici étaient trop simples pour modéliser correctement les phénomènes observés dans les langes naturelles.

Par la suite, nous remarquons l'apparition de travaux importants, en particulier ceux de Ronald Kaplan et de Martin Kay. Les années 70 voient aussi une recrudescence des travaux en syntaxe, et l'émergence de nouveaux formalismes de description grammaticale.

Aujourd'hui, le domaine du traitement du langage naturel est un domaine de recherche très actif. De nombreux logiciels, facilitant des tâches à l'Homme, commencent à

---

[15] - H. Bersini, *De l'intelligence humaine à l'intelligence artificielle*, Ellipses, Paris, 2006, p.56.

atteindre le grand public, elles sont là pour témoigner de l'importance des progrès accomplis mais aussi des projets qui restent à accomplir.

Plusieurs disciplines et plusieurs recherches se développaient à cette époque là. Chacune de ces disciplines, telles que l'intelligence artificielle, les statistiques… croise le chemin de la linguistique. Nous allons définir dans ce qui suit quelques disciplines importantes dans le domaine du traitement automatique du langage naturel (TALN).

I-2-Disciplines regroupées en TALN

I-2-1-Linguistique

La linguistique est une étude scientifique du langage humain. Elle a pour principal objet d'étude « *la langue* » étudiée « *en elle-même et pour elle-même* »[16] pour reprendre les termes de Ferdinand De Saussure, père fondateur de la linguistique.

Selon le dictionnaire Dubois « *on s'accorde généralement à reconnaitre que le statut de la linguistique comme étude scientifique du langage assuré par la publication en 1916 du Cours de linguistique générale de F. De Saussure. A partir de cette date, toute étude linguistique sera définie comme apparue avant ou après Saussure.* »[17]. La linguistique s'intéresse à étudier le langage, qui est une faculté propre à l'homme, dans tous ses états. Ainsi, en traitement automatique du langage naturel, la linguistique fournit des théories explicites du savoir linguistique.

I-2-2-Intelligence artificielle

Nous avons toujours associé le terme « intelligence », qui signifie l'ensemble du savoir et du savoir faire, à l'être-humain jusqu'au jour où nous assistons à la naissance d'une nouvelle discipline que nous appelons « intelligence artificielle ».

Nous associons la naissance de cette nouvelle discipline à la première traduction automatique dans les années 50. Ainsi, en 1956, un groupe d'informaticiens s'intéressent à ce domaine et se réunissent pour une conférence autour de ce thème. Aujourd'hui, l'intelligence artificielle est considérée comme une discipline scientifique importante dans le domaine de l'informatique et aussi en traitement automatique du langage naturel.

A la naissance de l'intelligence artificielle, les chercheurs se sont fixés comme objectif de remplacer l'humain par un ordinateur dans toutes ses tâches. Ainsi, Hugues Bersini affirme que « *l'intelligence humaine demeure la source d'inspiration première et sans doute*

---

[16] - F. De Saussure, Op.cit. p.376.

[17] - J. Dubois, Op.cit. p.300.

le graal ultime pour sa contrepartie informatique. Les premières définitions de l'intelligence artificielle firent de l'intelligence humaine le but à atteindre, sinon à dépasser »[18], par ce passage, nous comprenons que la définition de cette nouvelle discipline tourne autour da la branche qui a pour but d'automatiser les facultés propres à l'Homme. Par ailleurs, en traitement automatique du langage naturel, l'intelligence artificielle s'occupe de la représentation des connaissances et de leur utilisation.

I-2-3-Informatique

L'informatique a pris de plus en plus une place importante dans les sciences du langage. Par le terme lui-même « informatique » nous pouvons avoir une première définition. Ainsi, le mot « informatique » est composé des deux termes : « information » et « automatique ».

A l'origine, il désigne l'ensemble des activités liées à la conception et à l'emploi des ordinateurs pour traiter des informations. Il désigne aussi l'ensemble des sciences formelles qui ont pour objet d'étude la notion d'information et des procédés de traitement automatique de celle-ci.

L'informatique moderne commence avant la Seconde Guerre mondiale, lorsque le mathématicien Alain Turing pose les bases d'une théorisation de ce qu'est un ordinateur, avec son concept de *machine universelle de Turing*[19].

L'informatique joue un rôle important dans le domaine du traitement automatique du langage naturel. Elle permet d'optimiser des algorithmes et les programmes de traitement, mais aussi de développer des techniques formelles.

I-2-4- Linguistique informatique

La linguistique informatique est une sous branche de la linguistique et de l'informatique. Elle voit le jour par le croisement de ces deux disciplines. Avant toute tentative de la définir, il faut signaler, au départ, la différence entre ce que nous désignons par « linguistique informatique » et « informatique linguistique »

Ainsi, l'informatique linguistique et la linguistique informatique n'ont pas le même rapport et ne se réfèrent pas à la même linguistique. L'informatique linguistique s'est constituée autour d'une linguistique retreinte pour l'essentiel à la morphosyntaxe,

---

[18] - H. Bersini, De l'intelligence humaine à l'intelligence artificielle, Ed Ellipses, Paris, 2006, p.15.
[19] - H. Dreyfus, *Intelligence artificielle, mythes et limites*, Flammarion, Paris, 1984, p.5.

et « *produit toute une gamme d'analyseurs morphologiques et syntaxiques* »[20]. En revanche, la linguistique dite informatique utilise le savoir informatique pour les besoins de la linguistique. Nous la connaissons aussi sous le nom de *linguistique assistée par ordinateur*[21].

Par ailleurs, la linguistique informatique se divise en deux branches. La première répond au besoin de la linguistique théorique, elle a permis, par exemple, la mise au point de logiciels d'aide à la reconstruction des évolutions du lexique. La seconde répond, quant à elle, aux besoins de la linguistique appliquée ; dans le cas de la traduction automatique, par exemple.

Par ailleurs, le lien entre ces disciplines se résume en l'étude du langage naturel. Ainsi, en premier lieu, l'informatique joue un rôle important, car d'une part l'intelligence artificielle a eu une influence importante sur la linguistique en général ; et d'autre part la numérisation de nombreux textes et le développement des besoins sociaux liés à leur gestion et à leur traitement exige des solutions pratiques. En deuxième lieu, l'intelligence artificielle a souhaité modéliser la compréhension du langage, et influencé par la linguistique chomskyenne, elle a mis l'analyse syntaxique au centre de ses préoccupations. En dernier lieu, la relation entre l'informatique et la linguistique peut s'expliquer par le fait que l'analyse linguistique est un préalable au traitement informatique, elle permet une préanalyse du corpus, en fonction de la tâche.

I-3-Les domaines du traitement automatique des langues

I-3-1-Traitement de la parole

Le décodage acoustico-phonétique est considéré à la fin des années 70 comme le problème majeur de la reconnaissance. Ainsi, trois approches sont envisagées. Nous avons en premier lieu, l'approche fondée sur le traitement du signal et la reconnaissance de formes ; en deuxième lieu nous avons l'approche qui se fonde sur une catégorie de type intelligence artificielle ; quant à la dernière approche, toujours fructueuse aujourd'hui, est axée sur un point de vue probabiliste de décodage acoustico-phonétique.

I-3-2-Compréhension automatique des textes

Comprendre un texte, c'est en effet, par-delà le simple décodage du contenu littéral de ce qui est dit, c'est-à-dire être capable de relier les phrases entre elles de façon à

---

[20] - Sabah G, *L'intelligence artificielle et le langage*, Ed Hermès, 1989, p56.
[21] - Y.M, Visetti. Modèles connexionnistes et représentation structurées, Intellectica, 1990, p.10.

reconstruire un tout signifiant et cohérent, et être capable d'interpréter le message reçu par rapport à la situation et aux conditions d'énonciation.

Ainsi, l'élaboration de systèmes de compréhension automatique de textes écrits se heurte à deux problèmes liés, d'une part aux relations interphrastiques, et d'autre part au contexte. Or, ces difficultés n'empêchent pas que nous arrivons aujourd'hui à traiter des aspects limités du sens d'un texte, de manière à pouvoir l'indexer correctement dans une base de données. Par ailleurs, nous pouvons aller plus loin pour obtenir une compréhension plus profonde mais cela serait relatif à un domaine très précis.

I-3-3-Génération automatique des textes

Ce n'est qu'au début des années 1980 que le problème de génération de textes, dans un acte de communication donné est abordé. Dans ce cadre, le texte généré par la machine doit satisfaire des exigences : d'une part indiquer à l'utilisateur les informations qu'il désire, et d'autre part offrir une formulation de ces informations dans une langue correcte.

Ainsi, il s'en suit que le processus de génération comporte deux composantes : la première traite la question « quoi dire ? », c'est-à-dire il détermine le contenu informatif, c'est un système expert de raisonnement. Quant au second, il traite la question « comment le dire ? », autrement dit, il formule le contenu informatif dans une langue correcte, ce dernier est le module de génération linguistique.

I-3-4-Dialogue homme-machine

C'est une interaction avec la machine par le biais du langage ; un système de dialogue homme-machine comporte, outre un module d'analyse et un module de génération, un module de raisonnement spécifique. Raisonner suppose la capacité d'accéder à une base de connaissances quelconque par exemple les horaires des trains ou des avions.

Il existe d'autres domaines du TALN, mais nous allons nous contenter de ces exemples que nous venons de citer ci-dessus. Nous pouvons citer, entre autre, la correction orthographique, la gestion électronique de l'information, et la traduction automatique qui constitue un élément important dans notre thème de recherche, et que nous allons détailler dans le paragraphe suivant.

II-La traductique

II-1-La traduction

La traduction est une activité de la pensé humaine. Son but est de réaliser des ponts entre différentes langues, donc, entre différents modes de communication. Elle est une

sorte de lien entre deux moyens d'expression, deux imaginaires, deux cultures, voire deux mondes différents.

Nous avons commencé à utiliser le terme de traduction au sens de faire passer un ouvrage d'une langue à une autre. Ainsi, le Petit Robert définit le verbe traduire comme un : « verbe qui provient du latin traducere (1480) » ce dernier signifie « *faire passer* », c'est aussi « *faire que ce qui était énoncé dans une langue naturelle le soit dans une autre, en tendant à l'équivalence sémantique et expressive des deux énoncés* »[22].

Selon Eco, la traduction c'est « *tenter de comprendre comment, tout en sachant qu'on ne dit jamais la même chose, on peut dire presque la même chose… c'est dire presque la même chose dans une autre langue.* »[23]. Le mot *presque* nous révèle qu'Eco voulait dire qu'une traduction parfaite n'existe pas. Autrement dit, du fait que chaque langue ait sa propre vision du monde et sa propre structure, nous ne pouvons pas avoir un texte traduit identique au premier.

La traduction est l'opération qui consiste en un passage d'un texte en langue source Ls à un texte en langue cible Lc, sans toutefois que ne soit affecté le sens sémantique général. Autrement dit, traduire c'est énoncer dans une autre langue ce qui a été énoncé dans une langue source, tout en tenant compte d'un certain nombre de contraintes (comme le contexte, l'équivalence, la syntaxe…), afin de le rendre compréhensible par des personnes ne connaissant pas la langue source et n'ayant pas la même culture.

II-1-1-L'équivalence en traduction

La tâche de traduire n'est pas simple. Le traducteur, considéré comme le médiateur de la communication interlinguistique et interculturelle, se doit de rechercher une équivalence qui rend le texte d'arrivée fonctionnel dans la culture réceptrice. Nida affirme que « *la traduction consiste à produire dans la langue d'arrivée l'équivalent naturel le plus proche du message de la langue de départ, d'abord quant à la signification, puis quant au style* »[24]. Une de ces idées est la notion de symétrie qui existe dans la valeur des éléments des différentes langues. Nous pouvons conclure selon la citation de Nida que l'équivalence existerait à condition qu'elle ait la même valeur que l'original. Cependant, le fait que chaque

---

[22] - *Le nouveau petit Robert de la langue française*, Paris, 2oo8, 2592.
[23] - Eco Umberto, *Expériences de traduction*, Grasset, Paris, 2007, p.10.
[24] - Nida, In Eco Umberto. Op. Cit. p.38.

langue organise son espace sémantique de façon différente rend difficile toute possibilité de valeur égale entre les éléments des différentes langues.

II-1-2-Les différents types d'équivalence

II-1-2-1-L'équivalence linguistique

Nous appelons équivalence linguistique, ou appelée aussi équivalence formelle, le type d'équivalence qui présente une homogénéité sur le plan linguistique ente le texte original, c'est-à-dire celui de la langue source, et le texte traduit, ou celui de la langue cible. Nous parlons généralement d'équivalence linguistique quand la traduction est faite mot-à-mot, autrement dit quand nous traduisons les mots et non pas le sens global du texte.

II-1-2-2-L'équivalence paradigmatique

Ce type d'équivalence se caractérise par l'imposition d'une certaine correspondance, sur le plan grammatical, entre la traduction et le texte original. Il s'agit d'éléments grammaticaux pouvant être substituables sans que le sens de l'énoncé ne soit modifié. En français, en espagnol, par exemple, le temps verbal des lois est généralement le présent de l'indicatif. Ce temps peut être substitué par un autre si les usages et coutumes du système de la culture réceptrice l'exigent. L'équivalence paradigmatique désigne des rapports dits verticaux, par opposition aux rapports syntagmatiques, plutôt qualifiés de rapports horizontaux. En général, le traducteur obtient cette équivalence par le biais de la transposition.

II-1-2-3-Equivalence stylistique

En équivalence stylistique, les linguistes parlent d'une relation fonctionnelle entre les éléments stylistiques du texte de départ et du texte d'arrivée. Cette relation permet d'voir une identité expressive ou affective entre l'original et sa traduction, sans modifier le sens global de l'énoncé à traduire. Le style donné à la traduction peut varier selon le texte à traduire ; c'est-à-dire que ce soit un texte littéraire ou un texte pragmatique. En traduction juridique par exemple, le style du texte d'arrivée doit être conforme aux usages du système juridique de la culture réceptrice.

II-1-2-4-Equivalence sémantique

Nous parlerons d'équivalence sémantique lorsque le texte de la langue source et le texte de la langue cible ont le même contenu sémantique ou sémiotique (le signifié). L'équivalence sémantique se place sur le plan des mots et non du paragraphe ou du texte dans son ensemble. Elle suppose que le terme de départ et son équivalent en langue

d'arrivée partagent un champ sémantique identique. Dans ces cas, la traduction littérale s'avère un procédé de traduction approprié.

II-1-2-5-Equivalence formelle

Nous appelons aussi ce type d'équivalence par équivalence textuelle, syntagmatique ou encore équivalence structurelle. L'équivalence formelle est employée pour reproduire le plus littéralement possible le contenu et la forme du texte de départ de façon à ce que les récepteurs de la traduction reçoivent le même message, dans son contenu comme dans sa forme. Elle permet de s'identifier le plus fidèlement possible à la personne ou à l'objet dont il est question dans le texte de départ, afin de comprendre le mieux possible ses traits culturels : ses coutumes, sa façon de penser, et ses moyens d'expression. L'équivalence formelle n'est nulle autre que la traduction littérale. La traduction par équivalence formelle a comme caractéristique d'être une traduction orientée vers le texte de départ. En fait, c'est ce dernier qui sert de référence dans l'établissement des normes de fidélité. La traduction glosée est le parfait exemple d'une traduction rendue par équivalence formelle.

II-1-3-La notion d'interprétation

La notion d'interprétation est une notion qui se trouve dans plusieurs domaines tels que le théâtre, la littérature, et aussi la traduction. Ainsi, en traduction l'interprétation a pour fonction d'élucider le sens d'un texte ou d'un acte. Par ailleurs, la nécessité de l'interprétation tient en effet à ce qu'il n y a pas de réception immédiate du sens parce que le sens d'un texte quelconque est généralement ambigu. Des expressions de la langue nous laissent comprendre tout à fait le contraire de ce qui est dit en apparence.

Le rôle de l'interprète est de chercher à restituer fidèlement le sens de l'objet de son interprétation, sa tâche est difficile par le fait de la difficulté de rester fidèle au sens du texte.

Nous pouvons dire, donc, que l'interprétation vise à combler un déficit de sens et donner ainsi de la rationalité à ce qui n'en aurait pas suffisamment par lui-même. Ainsi, interpréter ne signifie pas « expliquer » comme le pense la plupart ; la différence entre ces deux mots est très claire, parce que nous expliquons un phénomène, scientifique par exemple ; mais nous interprétons un texte.

En outre, interpréter veut dire aussi élucider, c'est-à-dire mettre en évidence l'implicite. Or, cette opération est difficile quand nous passons d'une langue à une autre. Ainsi, en traduction, nous sommes face à l'interprétation du texte que nous voulons traduire mais en lui restant fidèle.

II-2-Premiers pas en traduction automatique

La *traductique* ou traduction automatique est le premier domaine du TALN, autrement dit, elle est l'application la plus ancienne du TALN. Elle est aussi l'un des objectifs historiques de l'intelligence artificielle.

La traduction automatique ou TA vise à traduire un texte d'une langue source à une langue cible automatiquement. Ainsi, les recherches en traduction automatique ont été entreprises au début des années 1950, dès l'apparition de l'ordinateur. La traduction automatique est une application qui prend une place de plus en plus importante dans la vie de tous les jours.

La première opération de traduction automatique a été faite dans les années 1950, nous avons traduit quelques phrases russes vers l'anglais. C'est en effet deux facteurs politiques, la recherche du dialogue et la compétition réelle entre nations, qui ont été déterminants pour le développement de la traduction automatique. Or, cette traduction était trop simplicité car nous ne traduisons que quelques phrases simples. Les premiers systèmes de traduction automatique reposaient sur une traduction mot-à-mot, mais cette traduction fut vouée à l'échec parce que nous savons que nous ne pouvons pas traduire sans comprendre.

Les chercheurs se posaient la question sur la possibilité qu'un ordinateur puisse traduire. Ainsi, dès les débuts de l'aventure de traduction automatique, il était possible de stocker dans un ordinateur un ensemble de mots, et par la suite même à stocker deux ensembles de mots : un ensemble de mots russes et un ensemble de mots anglais, puis indiquer à l'ordinateur quel mot anglais correspondait à quel mot russe, c'est-à-dire ils ont lié automatiquement les significations des mots anglais et des mots russes.

Or, la langue n'est jamais un ensemble de signes utilisés séparément, mais c'est un ensemble de signes qui se lient selon des règles pour former des phrases, puis des textes. Ainsi, la difficulté rencontrée consiste en la manière de rendre un ordinateur capable de dégager le sens d'une phrase à partir de la forme de cette phrase même. Pour s'y faire,

l'ordinateur doit être capable de connaitre, outres les deux ensembles de signes et leurs significations, la grammaire des deux langues, donc le russe et l'anglais. Telle est l'une des deux raisons pour lesquelles les premiers pas en traduction automatique se trouvaient dans une impasse.

Par ailleurs, la langue n'est pas un tout, un système qui peut être étudié séparément du monde qui nous entoure. Ainsi, même si nous connaissons parfaitement la grammaire de nos langues, nous avons besoin d'une masse d'informations sur le monde pour pouvoir dégager le sens d'une expression donnée. En revanche, toutes ces difficultés et les craintes, d'une impossibilité de douer un ordinateur de toutes les connaissances nécessaires pour comprendre une langue, mènent à un arrêt de financement de tout projet de traduction automatique.

Par la suite, d'une part, la Communauté européenne a été créée et fourni de nouveaux financements pour la traduction automatique. Et d'autre part, l'ère de la traduction avait commencé. Ainsi, nous assistons à la création de logiciel de traduction même si la qualité de la traduction était loin d'être bonne.

Au fil du temps, les machines à traduire ont beaucoup évolué. Des systèmes très avancés sont apparus et donnent des résultats encourageants. Ainsi, au XX siècle, plusieurs phénomènes apparaissent et entrainent le développement des outils de traduction automatique. Nous citons par exemple la culture de l'écran qui s'oppose à celle de l'écrit. La communication tend à ne plus passer par l'écrit mais par l'électronique.

En outre, la traduction automatique comprend différents intermédiaires : nous avons en premier lieu, la traduction automatique assistée par l'homme (TAAH), qualifiée aussi de traduction assistée par ordinateur (TAO) ou tout simplement de traduction automatique assistée (TAA). Cette traduction ne peut pas s'opérer sans l'intervention de l'homme. En second lieu, nous avons la traduction entièrement automatique de haute qualité (TAHQ).

II-2-1-Les systèmes de traduction automatique

La traduction automatique est marquée par deux étapes. Ainsi, la première étape se marque par l'apparition de systèmes directs, et la seconde par les systèmes indirects.

II-2-1-1-Les systèmes directs

Dans ses débuts, la traduction automatique était caractérisée par une traduction mot-à-mot. Ainsi, pour traduire son texte, le système de traduction fait une analyse morphologique puis passe directement à la consultation de dictionnaires bilingues

enregistrés dans leur mémoire. Par la suite, le système passe à une autre étape où il essaye de ranger les mots selon des règles locales. Enfin, le système nous donne le texte traduit dans la langue cible. Beaucoup de systèmes ont été crées dans cette optique, nous pouvons citer par exemple le système *Systran* qui a été crée dans les années 60.

Or, ces systèmes sont trop simplifiés et donnent des traductions erronées. Dans des cas plus simples, les systèmes directs peuvent être utiles, comme dans le cas du système canadien TAUM-Météo qui traduit des bulletins météo régionaux du Canada de l'anglais vers le français, avec un vocabulaire très limité (moins de 1000 mots).

II-2-1-2-Les systèmes de transfert

L'échec des systèmes directs a poussé les chercheurs en TA à penser à d'autres systèmes plus avancés et plus utiles. Ainsi, durant les années 70, les systèmes de transfert voient le jour.

Les systèmes de transfert tel Metal possèdent trois modules. Le premier consiste à l'analyse en langue source ; la deuxième établit la génération en langue cible ; quant au dernier, c'est un module de transfert qui établit les équivalences entres les deux langues, c'est-à-dire entre la langue source et la langue cible. Par ailleurs, ces systèmes sont caractérisés par des dictionnaires bilingues et des règles d'équivalences lexicales selon le contexte où le mot de l'expression apparait.

L'analyse linguistique du texte en langue source permet d'obtenir une représentation intermédiaire indépendante des langues sources et cibles. A cette fin le système utilise une interlangue en faisant appel soit à des primitives sémantiques ou logiques soit à un modèle du domaine, voire à la description des connaissances du monde. Par ailleurs, c'est à ce niveau que la génération de texte en langue cible intervient.

A cette époque, les systèmes de transfert étaient considérés comme les systèmes les plus avancés en traduction automatique, c'est-à-dire par rapport aux systèmes directes, parce qu'ils donnent des résultats plus satisfaisants en matière de traduction.

II-2-2-Mémoire de traduction

Une mémoire de traduction est une base de données qui sert à garder un grand nombre de phrases, de textes et leurs traductions. Par ailleurs, une mémoire de traduction permet d'utiliser des listes terminologiques, grâce auxquelles l'ordinateur recherche ou ignore des termes spécifiques lors de l'établissement des parties de texte à traduire. En outre, ces listes peuvent être mises à jour au fur et à mesure.

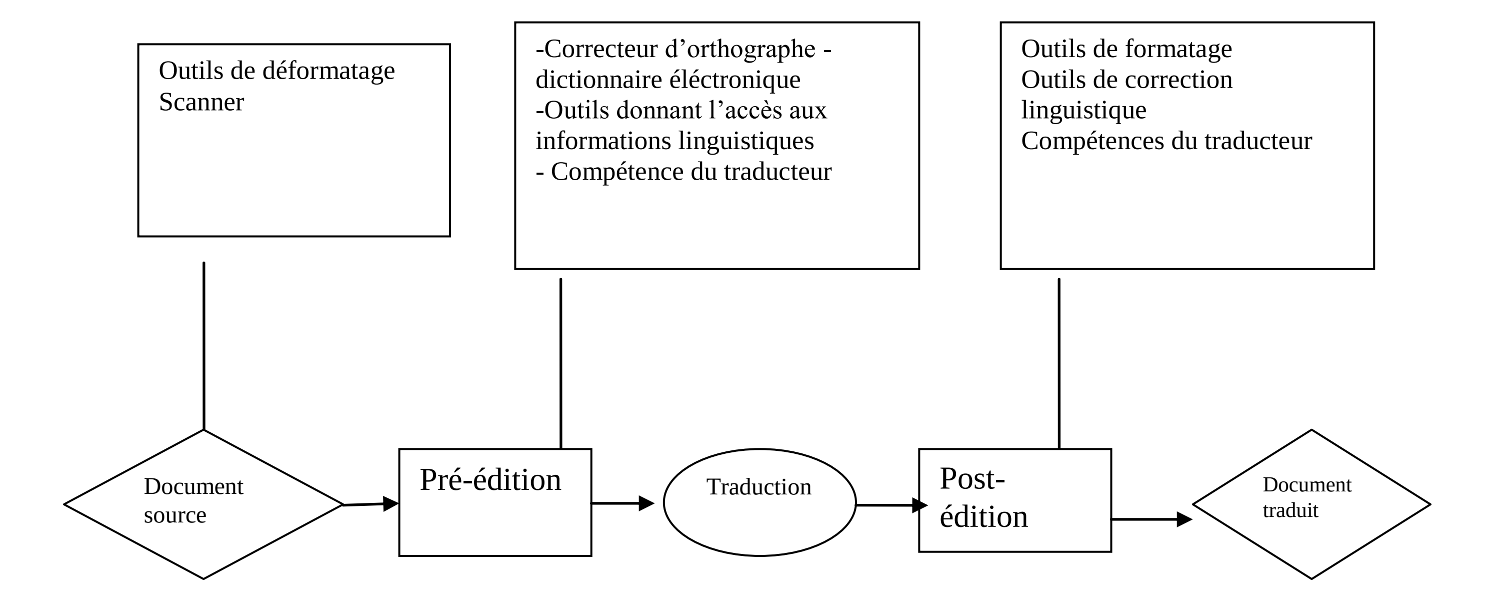
La mémoire de traduction permet non seulement de réutiliser la terminologie déjà fixée mais aussi de tirer des segments entiers de textes déjà traduits par le biais d'algorithmes de correspondance partielle évolués, afin d'assurer l'homogénéité de l'ensemble.

La mémoire de traduction est généralement composée des éléments suivants :

- Un éditeur de traduction, qui protège le format, c'est-à-dire la mise en forme du texte source ;
- Des fonctions de localisation des textes inclus dans des programmes et de prise en charge des informations en ligne et de tout autre texte crée sur un système ;
- Un outil de terminologie qui permet de gérer les dictionnaires créés par le traducteur ou importés par lui dans son système ;
- Un moyen d'analyse automatique de tout nouveau texte qui permet la création ou la mise à jour de listes de terminologie à utiliser postérieurement en phase de traduction ;
- Un outil statistique qui indique, entre autres, le nombre de mots traduits et à traduire et la longueur des phrases.

Nous vous proposons un schéma du processus de traduction automatique extrait de l'ouvrage *la traductique*[25]

[25] - P. Bouillon, *La traductique*, Ed Les presses universitaires de Montréal AUPELF URUF, Montréal, 1993, p.57.



Processus de traduction avec outil de traduction automatique

Nous concluons ce premier chapitre en synthétisant l'idée que, de nos jours, le traitement automatique du langage naturel TALN est un domaine important en linguistique et en informatique. Il est également un domaine vaste dont les tâches sont difficiles à réaliser mais qui restent possibles. Par ailleurs, la traduction automatique TA témoigne des grands pas effectués en traitement automatique des langues, et ceci par les divers moteurs de traduction qui existent sur le Web.

Chapitre II

Les théories linguistiques au service de la polysémie

Après avoir exposé le domaine du traitement automatique du langage naturel, nous passerons maintenant à une présentation de la polysémie, qui constitue un élément central de notre étude.

L'étude de la polysémie est très vaste et beaucoup de linguistes et sémanticiens se sont intéressés à l'étude de celle-ci. Nous essayerons dans ce deuxième chapitre de présenter les différentes notions clefs dans l'étude de la polysémie afin de mieux comprendre la notion elle-même de polysémie.

I-Le signe linguistique

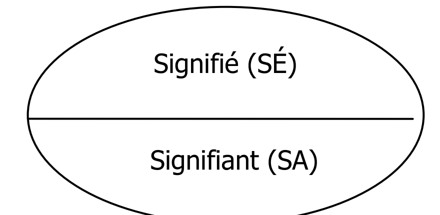
I-1-Le signe linguistique selon Ferdinand De Saussure

Le *signe linguistique*[26] consiste l'un des éléments les plus importants à évoquer dans une étude sémantique. Ainsi, nous ne pouvons pas faire une étude sur la polysémie sans aborder et définir le *signe linguistique*.

Depuis Ferdinand De Saussure (père fondateur de la linguistique moderne), le *signe linguistique* constitue un point important dans les études linguistiques. De maintes définitions ont été données par d'autres linguistes.

Dans notre étude, nous essayerons de présenter quelques définitions afin de mieux comprendre ce concept. Comme dans toutes les études en linguistique, nous commencerons par la définition du père fondateur de la linguistique moderne Ferdinand De Saussure. Ensuite, nous allons essayer de définir le *signe linguistique* selon d'autres linguistes.

Pour Saussure, le *signe linguistique* est un élément fondamental de la linguistique. La langue, étant l'objet d'étude de cette dernière, se définit comme « *un système de signes* »[27]. Ainsi, pour définir le *signe linguistique*, Saussure le compare au recto et au verso d'une feuille. Le recto étant le *signifié*, c'est-à-dire son concept (représentation mentale d'une idée ou d'une chose), et le verso est le *signifiant*, c'est-à-dire son image acoustique (représentation mentale d'une suite sonore). Autrement dit « *le signe linguistique est une combinaison, une association d'un signifié et d'un signifiant* »[28].



Le signe linguistique selon Ferdinand De Saussure

[26] - F. De Saussure, *Cours de linguistique générale*, Payot, Lausanne, 1995, p.109.
[27] - Ibidem
[28] - Ibidem

En outre, Saussure ajoute que ces deux aspects donnent au *signe linguistique* deux caractères fondamentaux : arbitraire et linéaire. Par arbitraire, nous entendons que la relation qui unit le *signifié* et le *signifiant* n'est pas logique. En d'autres termes, il n y a pas de relation logique entre le *signifiant* et le *signifié*. Mais ce lien est conventionnel, c'est-à-dire que les membres de la même *communauté linguistique*[29] se mettent d'accord pour nommer tel objet de telle manière. Sans que ce lien soit naturel : la représentation de la suite sonore /a r b r/ n'a aucune attache avec l'idée même ou à la notion de « arbre » dans la réalité. Il s'agit, donc, d'une entité psychique et sociale qui relève du système de la langue et non de la libre réalisation individuelle.

Quant au caractère linéaire du *signe linguistique*, il est dû à la nature auditive du *signifiant*, qui se déroule dans une dimension temporelle et cela impose une organisation particulière au système de signes qu'est la langue.

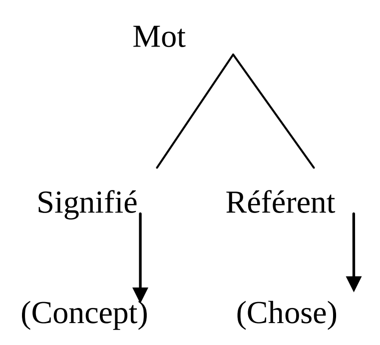
Après avoir introduit les notions du *signifiant* et du *signifié*, Saussure explique ce qu'il appelle par *valeur du signe linguistique*.[30] Ce dernier n'a de sens que par rapport aux autres signes à l'intérieure du système linguistique et c'est ce qu'il lui donne sa valeur. De ce fait, on appelle *valeur*, le sens d'une unité linguistique considéré comme le produit des relations que cette unité entretient avec les autres unités du système. Ferdinand De Saussure illustre cette notion par la comparaison de la langue avec le jeu d'échecs : n'importe quelle pièce de ce jeu, ou son support, ou le matériau dans lequel ils sont faits sont parfaitement indifférents au déroulement de la partie, pourvu que les joueurs se soient accordés sur la « valeur » à donner à l'élément. De même dans la langue, un élément ne s'identifie que par la valeur dont l'a anéanti la commune des usagers, à savoir le champ de fonctionnement qu'il recouvre à l'intérieur du système.

I-2-D'autres définitions du signe linguistique

Contrairement à Ferdinand De Saussure, d'autres linguistes ont préféré inclure le référent, c'est-à-dire l'entité extralinguistique à laquelle réfère le signe linguistique, dans leurs représentations du signe linguistique. Saussure exclut le référent de son signe car « *tout se passe entre l'image auditive et le concept, dans les limites du mot, considéré comme un domaine fermé, existant pour lui-même*. »[31]

[29] - Marie louise Moreau, *Sociolinguistique : concepts de base*, Ed Mardaga, Paris, 1997, p.88.
[30] - F. De Saussure, Op.cit. p.185.
[31] - Ibid, pp.158-159.

Tamba-Mecz propose, par exemple, un triangle sémantique (représentatif du signe linguistique) pour mieux apercevoir le « *rôle du pivot formel* »[32] que joue le mot lexical dans la mise en place et la régulation du double processus de dénomination-signification, qui constitue une pièce maitresse des systèmes linguistiques. Son triangle sémantique « *décroche* », comme elle dit, « *d'une part le signifié du concept* » et, d'autre part, « *le référent de la chose* », le « mot » a alors le double rôle de signe et de dénomination[33].



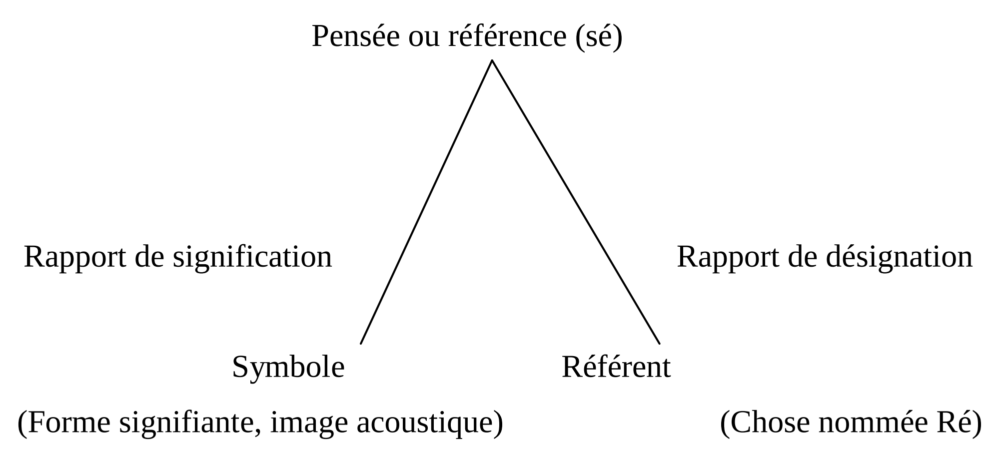
Le signe linguistique selon Tamba-Mecz

Le « mot » est ici la forme lexicale de l'unité du signifié et du référent, ceux-ci sont présentés « *au moyen d'une seule et même forme* »[34]. Le mot est à la fois le signifiant du signifié et le nom du référent. Le signifié relève donc de l'univers conceptuel, tandis que le référent représente toute la chose ou phénomène extralinguistique, au moyen de la dénomination.

Alors que Saussure considère le sens avant tout comme le résumé de la valeur, C.K. Ogden et I.A. Richards, tentent de réintroduire la chose nommée extralinguistique qui peut être réelle ou imaginaire dans la définition du signe. D'après eux « *nous ne pouvons parler de signifié que si les mots peuvent être tournés vers, référés à, en eux- mêmes ils ne signifient rien* »[35]. Leur célèbre triangle sémiotique a le mérite de mettre en évidence le rapport de signification que unit un (sa) et un (sé), et celui de désignation qui concerne le signe dans la liaison à l'objet.

---

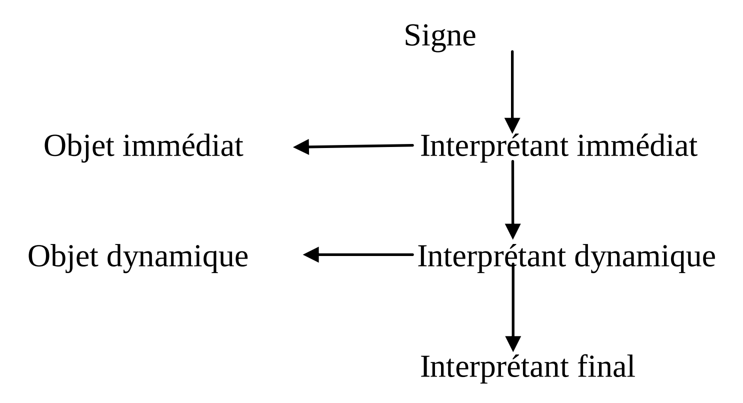
[32] - Tamba-Mecsz, I, *Le sens figuré*, Presses universitaires de France, 1981, Paris, p75.
[33] - Tamba-Mecsz , ibid, p.75.
[34] - Ibidem
[35] - C.K. Odgen, et I.A. Richard, *The meaning of meaning*, In jean Michel adam, *Linguistique et discours littéraire : théorie et pratique des textes*. Larousse université, Canada, 1976.



Le signe linguistique selon Ogden et Richard

Peirce distingue un *objet immédiat*[36], c'est-à-dire un référent au sens strict, fixé, sans lequel le signe n'existerait pas, mais qui ne recouvre pas toutes les possibilités existantes et un *objet dynamique,* référent plus large, qui comprend ce que le signe ne peut pas directement exprimer, mais ne peut qu'indiquer, et que le récepteur doit interpréter grâce à son expérience.

De même, il distingue un *interprétant immédiat*, c'est-à-dire un sens probable, susceptible de venir spontanément à l'esprit de n'importe quel récepteur qui connait le code, un *interprétant dynamique*, c'est-à-dire le sens particulier formé dans l'esprit d'un locuteur particulier à chaque instance de réception, et aussi d'un *interprétant final,* sens sur lequel tous les récepteurs peuvent s'accorder



Le signe linguistique selon C.S. Peirce

Ce modèle comprend certains éléments manquants dans ceux que nous avons vus précédemment. La distinction entre l'immédiat et le dynamique permet de séparer ce qui est reconnu, décodé de façon mécanique, et l'interprétation effective par un récepteur; par exemple, là où les définitions d'un mot telles qu'elles sont consignées dans un dictionnaire et le sens d'un mot en contexte.

---

[36] -Charles Sanders Peirce. *Collected Papers*. 8 vol. Cambridge, Harvard U.P., 1931-1958.
Writings of Charles S. Pierce: A Chronological Edition. Bloomington, Indiana U. P., 1982-1989.

Nous pouvons, ainsi, saisir la nuance fondamentale entre *dénotation*, le sens qui est dans le signe, nécessairement invoqué par le signifiant, et la *connotation*, sens hors du signe, mais sanctionné par la culture, c'est-à-dire qui ne résulte pas d'une interprétation purement personnelle. Ainsi, la dénotation de «chômage» nous est donnée par la définition du dictionnaire («Arrêt de travail consécutif à un licenciement), mais chez les Français de la fin du XXe siècle et du début du XXIe, ce mot prend une connotation liée à la prééminence du chômage en tant que problème social en France, les troubles psychologiques qu'il entraîne, la politique des divers gouvernements qui ont cherché à le réduire.

II-Sens, signification, référent et référence

II-1-Définition du sens

La définition des notions de « sens », « signification », « référent », et « référence » apparait comme problématique. Ainsi, nombreux sont les linguistes et les sémanticiens qui ont consacré beaucoup de réflexions et beaucoup d'études pour les définir. Les conceptions qui existent, varient souvent d'un auteur à un autre, et elles sont parfois mêmes très différentes les unes des autres.

Nous allons essayer de présenter, dans ce qui suit, quelques définitions afin de mieux comprendre ces notions. La compréhension de toutes ces notions nous aidera à bien cerner la notion de « *polysémie* »[37].

La notion de « sens » a toujours fasciné les chercheurs et les linguistes. Elle est le pivot de toutes les études en sémantique. Dans son ouvrage, *Pour aborder la linguistique*, Roland Eluard résume les définitions des linguistes en deux types de réponses. La première est celle « *où le sens découle du jeu des différences à l'intérieur d'un système de signes* »[38], et la seconde est celle « *où le sens découle d'un rapport de référence à l'univers extra-linguistique* »[39].

Nous essayerons dans ce qui suit de présenter quelques définitions du sens selon chaque linguiste. Nous aborderons, en premier lieu, les définitions liées au système de signes ; puis nous entamerons, en second lieu, celles liées à des critères extralinguistiques. Ainsi, pour Saussure, c'est la première définition qui convient puisque le signe est présenté comme l'association arbitraire d'un *signifié* et d'un *signifiant* à l'intérieure d'un système de signes. Donc, Ferdinand De Saussure enferme le sens dans le signifié.

[37] - M. Bréal, *Essai de sémantuque : sciences des significations*, Ed Hachette, Paris, 1897, p.30.
[38] - Roland Eluard, *Pour aborder la linguistique*, Ed ESF, Paris, 1993, p.135.
[39] -Ibid.

Dans le même ordre d'idées, Harris lie la définition du sens à la distribution de l'unité linguistique c'est-à-dire sa position dans le système linguistique par rapport aux autres unités. Il affirme que « *... tout morphème différent d'un autre dans sa distribution doit aussi différer de lui dans sa valeur sémantique.* »[40]. Par valeur sémantique, Harris entend le sens d'une unité définie par les positions relatives de cette unité à l'intérieur du système linguistique.

Contrairement à Saussure, Emile Benveniste distingue, dans son ouvrage *Problèmes de linguistique générale*[41], le sens et le signifié. Pour lui, le langage fonctionne à deux niveaux : celui des signes et celui de la phrase qui est la plus petite unité sémantique. Le problème du sens ne se pose donc pas au niveau du signe mais à celui de la phrase. C'est à ce niveau seulement que le système de signes prend du sens. Et ce sens découle d'un rapport de référence à l'univers extra-linguistique, la phrase étant définie comme un segment prédicatif, c'est-à-dire disant quelque chose à propos de quelque chose.

De leur côté, les béhavioristes renvoient l'étude du sens à une psychologie du comportement. L. Bloomfield définit le sens d'une unité comme la somme des situations où elle apparait comme stimulus et des comportements-réponses que ce stimulus entraine de la part de l'interlocuteur, c'est-à-dire que le sens d'une unité linguistique se construit et chez le locuteur et chez l'interlocuteur selon le comportement et l'interprétation de ces derniers dans différentes situations de communication où cette unité apparait.

Wagner, quant à lui, définit le sens comme suit : « *un sens se forme, par interprétation, à partir de données objectives, fournies par la situation, et à partir de données linguistiques, à savoir les ensembles lexicaux, déjà constitués dans lesquels s'insère tel signe nouveau* »[42]. Par ce passage, nous comprenons que pour Wagner le sens se construit sur deux niveaux : un niveau linguistique, c'est-à-dire dans le système même de la langue, et ceci par la relation des unités dans le *système de signes* ; et un niveau extralinguistique, c'est-à-dire en se référent à une réalité à l'extérieur du système, une réalité du monde qui nous entoure. Nous constatons, donc, que Wagner regroupe les deux types de définitions déjà présentées ci-dessus.

II-2- Sens et signification

La distinction entre les deux notions de sens et de signification n'est pas toujours claire alors que nous les opposons parfois. Nous avons défini la notion de sens et nous essayerons maintenant d'expliquer celle de signification en distinguant entre les deux.

---

[40] - R. Harris, *Language, Saussure and Wittgenstein. How to play games with words*, Ed Routledge, London, 1988, p.56.

[41] - Emile Benveniste*, Problèmes de linguistique générale 2*, Ed Gallimard, Paris, 1974. P.60.

[42] - R.L. Wagner et J. Pinchon, *La grammaire du français classique et moderne*, Hachette, 1962.

Tomba Mecz distingue entre le sens et la signification, il affirme que nous opposons « *… ainsi la signification (parfois appelée aussi dénotation) ou rapport entre mot et concept de chose ou chose au sens, ou ensemble de valeurs qui fixent la position respective de chaque terme à l'intérieur d'un réseau relationnel.* »[43]. Nous pouvons distinguer, par ce passage, entre signification et sens. La signification est le sens propre d'une unité linguistique, c'est-à-dire le sens approprié à cette unité conventionnellement avec la chose même dans la réalité extralinguistique. Autrement dit, c'est le sens dénoté qui lie l'unité linguistique avec le référent. Quant au sens, c'est la somme des valeurs qu'une unité puisse regrouper à chaque fois qu'elle change de position par rapport aux autres unités à l'intérieur du système de signes et à chaque réapparition dans un nouveau contexte.

George Kleiber définit la signification comme « *étant… la même pour l'émetteur que pour le destinataire de l'énoncé, c'est-à-dire qu'elle est objectif dans le sens d'être* »[44]. Nous pouvons déduire par la définition de Kleiber que la signification est objective quant au sens, il est subjectif. En outre, nous pouvons parler d'une seule signification et de plusieurs sens.

II-3- Sens, référent et référence

Ces deux notions de sens et de référence posent également plusieurs interrogations. Si nous utilisons ces deux notions cela signifie qu'elles sont différentes, et nous pouvons s'interroger sur la relation entre ces deux éléments nécessaires à la compréhension et au fonctionnement de la langue.

Pour Frege « *la référence n'est pas à inclure dans le sens* »[45], il donne l'exemple de « *L'étoile du soir* » et « *L'étoile du matin* ». Ces deux expressions ont bien des sens différents tout en partageant la même référence « *Planète Vénus* ». Nous concluons par cet exemple que le sens du mot n'est pas du tout son référent ou référence. Le référent est ce que dénote l'unité linguistique utilisée tandis que le sens est exprimé par ce qui est connoté à ce référent.

De son côté Kleiber considère que « *le sens d'une expression linguistique est constitué par des traits auxquels doit satisfaire une entité pour être désignée par cette expression linguistique, c'est-à-dire pour être son référent.* »[46], c'est-à-dire que chaque sens doit référer à quelque chose qu'on appelle un référent. Ainsi la relation qui unit une expression linguistique au référent est communément appelée référence.

---

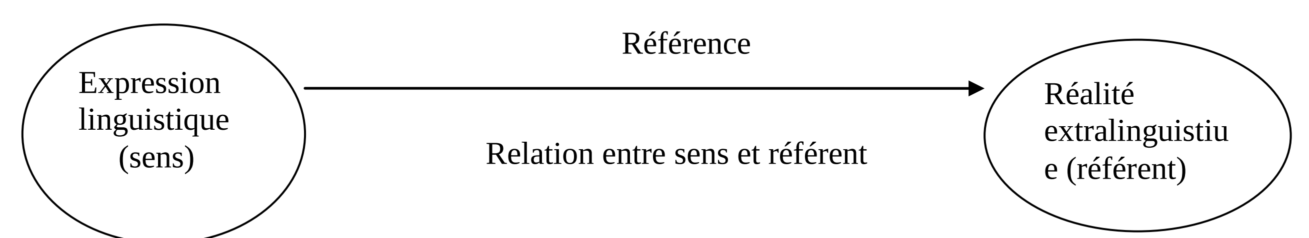
[43] - Tamba-Mecsz I, Op.cit. p.90.

[44] -G. Kleiber, *Polysémie et référence : la polysémie, un phénomène pragmatique ?*, Cahier de lexicologie n 44, 1984, p.90.

[45] - G. Frege, Sens et Référence, revue Zeitschrift für Philosophie und philosophische Kritik, 1892.

[46]- G. Kleiber, Op.Cit, p.110.

Nous pouvons résumer ces trois notions de sens, référence, et référent par le schéma suivant



III- La polysémie : une première définition

Après avoir défini quelques notions de base qui nous permettront de mieux comprendre la notion de la polysémie, nous allons passer maintenant à une définition de cette notion centrale autour de laquelle tourne notre problématique.

La notion de polysémie a toujours été problématique et elle l'est toujours dans toutes les études en linguistique, plus précisément en sémantique parce que c'est elle qui s'intéresse à l'étude du sens. Et aujourd'hui, elle est un problème central dans les recherches en traitement automatique des langues. En revanche, si nous voulons définir la polysémie, nous devons parler de plusieurs définitions et non pas d'une seule parce que chaque linguiste la définit différemment. Nous allons, donc, essayer de présenter quelques définitions dans le but de mettre au clair cette notion.

Pour se faire, nous allons commencer par présenter la notion de polysémie en allant de la définition la plus simple à une définition plus complexe. Ainsi, dans le dictionnaire Le Petit Larousse, la polysémie est définie comme « *la propriété d'un mot qui présente plusieurs sens.* »[47]. En reprenant presque la même formulation mais en s'étalant dans la définition, le dictionnaire Larousse de *linguistique et des sciences du langage*, Dubois définit la polysémie comme « *la propriété d'un signe linguistique qui a plusieurs sens.* »[48], c'est-à-dire qu'un même signifiant peut présenter plusieurs signifiés.

Par cette définition de la polysémie, nous remarquons que définir cette notion est une tâche difficile à accomplir à cause des termes qui la composent. Nous avons essayé, dans le chapitre précédent, de définir et de mieux comprendre la notion de sens et nous avons vu que cette notion est particulièrement délicate et fait partie des autres termes linguistiques qui ne se laissent pas facilement définir.

Ainsi, la première définition connue de la polysémie a été donnée par Bréal, selon lui « *le sens nouveau, quel qu'il soit ne met pas fin à l'ancien. Ils existent tous les deux l'un à*

---

[47]- *Dictionnaire Le Petit Larousse illustré*, Larousse, paris, 1991, p.259.

[48] - J. Dubois, *Dictionnaire de linguistique et des sciences du langage*, Larousse, Paris, 1994, p.150.

côté de l'autre. Le même terme peut s'employer tour à tour au sens propre ou au sens métaphorique, au sens restreint ou au sens étendu, au sens abstrait ou au sens concret… A mesure qu'une signification nouvelle est donnée au mot, il a l'air de se multiplier et de produire des exemplaires nouveaux, semblables de forme, mais différents de valeur. Nous appellerons ce phénomène de multiplication la polysémie. »[49]. Par ce passage nous comprenons que Bréal associe la notion de polysémie à un « *phénomène de multiplication* », c'est-à-dire que dans la polysémie il y a une sorte de production de sens nouveau par une même unité. Autrement dit, la polysémie peut être perçue comme la capacité des unités de prendre un sens nouveau.

De ce point de vue, nous pouvons considérer la polysémie comme un potentiel de la langue, celui de faire apparaitre des significations nouvelles pour des unités qui ne les offraient pas précédemment.

En outre, Roland Eluard dans son ouvrage *Pour aborder la linguistique* affirme qu' « *il y a polysémie quand plusieurs signifiés renvoient à un signifiant unique en une sorte de nébuleuse de significations.* »[50], c'est-à-dire qu'une même unité linguistique présente plusieurs sens. Nous présenterons le schéma de Roland Eluerd où il présente la notion de polysémie :

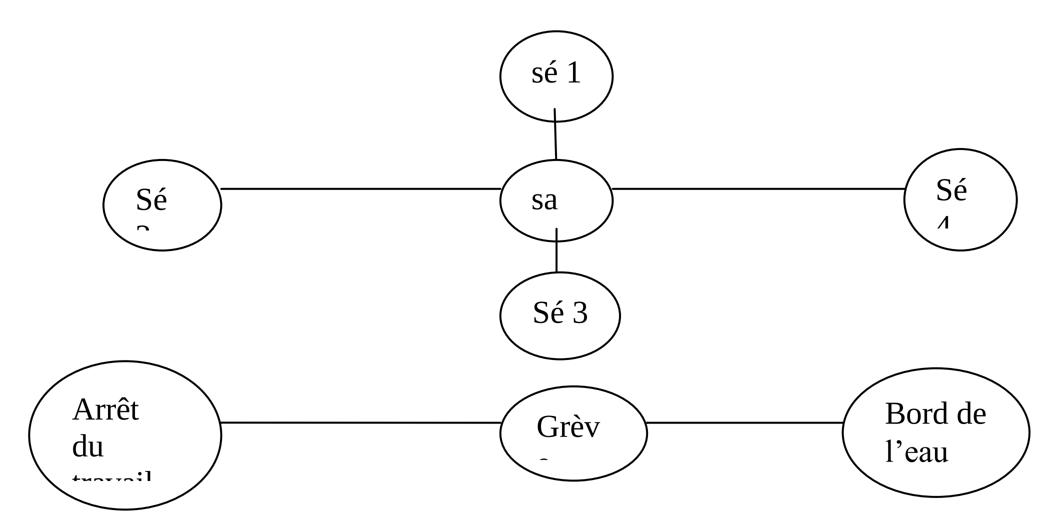


Schéma synthétisant la définition de la polysémie selon Roland Eluerd

Pour d'autres linguistes, il y a polysémie lorsqu'il y a « *existence de sens différents pour un même signe, mais des sens perçus comme reliés d'une manière ou d'une autre* ». Par ce passage de Cadiot et Habert nous ajoutons à la définition de la polysémie le fait que les sens d'une unité polysémique sont forcément liés à un premier sens.

[49] - M. Bréal, Op.cit. p200.
[50] - R. Eluerd, *Pour aborder la linguistique*, ED ESF. Paris, 1993, p. 139.

En revanche, Victorri et Fush affirment que définir la notion de polysémie n'est pas une tâche aisée. Pour eux *« la polysémie, comme beaucoup d'autres termes linguistiques, est une notion qui se laisse très facilement appréhender de manière intuitive, mais qui se révèle beaucoup rétive à une définition rigoureuse. On peut aisément se convaincre de la réalité et de l'importance de ce phénomène, en donner de nombreux exemples, et en décrire les principaux traits. Mais ces descriptions restent informelles, et dès que l'on cherche à dépasser ce niveau superficiel, pour tenter de caractériser exactement cette notion, d'en déterminer les limites précises, d'en donner des critères définitoires, on se heurte à la nécessité de construire un cadre théorique global avec son lot de définitions axiomatiques et de postulats toujours discutables, qui nous éloignent de la relation immédiate que nous entretenons avec la langue, et qui peuvent sembler d'une complexité démesurée par rapport à l'objet étudié.* »

En outre, la langue est en perpétuelle évolution et les unités qui la composent en sont les éléments les plus perméables. Toutefois, nous constatons l'apparition de nouvelles acceptions. Ainsi, la définition de la polysémie passe inévitablement par la reconnaissance de cette mouvance incessante et par l'acception de fait qu'aucune étude systématique ne demeurera pérenne.

Néanmoins, des linguistes comme Bianchi renvoient la multiplication des sens d'une unité polysémique au nombre de situations où cette unité apparait. Autrement dit, le nombre de sens qu'une unité linguistique peut avoir revient au nombre de contexte dont lequel cette unité est utilisée. Pour étayer nos dires, nous citons le passage de Bianchi où il affirme que « *le sens d'un mot peut varier de façon définie, tout comme le peut le* nombre de ses contextes *en emploi* »[51]

III-1- Définition de la polysémie selon la sémantique puissancielle

La théorie de *la sémantique puissancielle*[52] a été introduite par Guillaume. Elle consiste dans le fait de déterminer dans la polysémie un élément mental, résidant à la base de quelques représentations langagières existant réellement dans l'usage du langage. Pour Guillaume, nous définissons l'unité polysémique par le fait que dans celle-ci, nous trouvons ce que Gustave Guillaume appelle signifié de puissance et signifiés d'effet. Ainsi, Guillaume place le signifié de puissance dans la langue et les signifiés d'effet dans la parole. Autrement dit, le signifié de puissance est la première signification d'une unité linguistique, quant aux

---

[51] - CL. Bianchi, *La flexibilité sémantique : une approche critique, Langue française*, Larousse, Paris, p.100.

[52] - G. Guillaume, Principes *de linguistique théorique*, Ed Klincksieck, Paris, 1983, p.34.

signifiés d'effet se sont ses significations qui se rajoutent au premier dans l'usage de la langue, c'est-à-dire la parole.

Les notions de signifiés de puissance et de signifiés d'effet constituent les deux notions clefs de la sémantique puissancielle. Ainsi, Picoche explique aussi ces deux notions en affirmant que « *le signifié de puissance, réalité inconsciente, de l'ordre du virtuel, devient actuel par l'effet du discours. Capable d'un seul type d'actualisation, le signe est un monosème ; capable de divers types d'actualisation, le signe est un polysème* »[53].

Par ailleurs, Douay et Roulland ajoutent encore que « *c'est au plan de la puissance que se construisent les formes de langues, c'est là qu'elles prennent de par leur position en système leur valeur fondamentale, valeur qui se montre permissive à l'endroit d'un nombre infini de valeurs d'emploi ou effets de sens en discours* »[54]. Nous comprenons par les deux passages que nous avons cités, que la sémantique puissancielle considère le sens comme quelque chose d'unitaire qui se manifeste en discours au moyen d'acceptions différentes et ces sens forment une unité polysémique.

III-2- Polysémie vs homonymie

Nous avons vu ci-dessus ce qu'est la polysémie. Nous avons dit que cette notion sert à désigner la faculté d'un terme à prendre plusieurs sens. En outre, la notion de polysémie a été définie aussi en rapport avec l'homonymie. La différence entre ces deux notions a été beaucoup étudiée. Ainsi, nous allons essayer dans cette partie de distinguer ce qu'est la polysémie et ce qu'est l'homonymie.

Les linguistes tentent de distinguer la polysémie qui est la capacité d'un mot de prendre plusieurs sens, de l'homonymie qui est le fait pour une unité lexicale de constituer plusieurs mots. Donc, il s'agit dans la polysémie comme dans l'homonymie d'une multiplication du sens. Or, nous parlons dans la polysémie d'une seule unité lexicale et dans l'homonymie de plusieurs unités. En outre, la polysémie diffère de l'homonymie par le fait qu'elle présente aussi, dans les usages des termes, de fortes similarités sémantiques, que nous pouvons qualifier d'une sorte de relation de parenté.

Riegel, Pellaat et Rioul écrivent à ce propos que la « *polysémie s'oppose à l'homonymie lorsqu'il s'agit de décider… si une forme linguistique manifestant des contenus différents doit être traitée comme un seul ou plusieurs mots.* ».[55] C'est-à-dire, comme

---

[53] - J. Picoche, *Polysémie n'est pas ambiguité*, Cahiers de praxématique, Université Paul Valéry, Montpelier, 1989, p.8.

[54] - G. Douay et D Roulland, *Les mots de Gustave Guillaume : vocabulaire technique de la psychomécanique du langage*, Presses Universitaires Rennes II, Rennes, 1990, p. 155.

[55] - M. Riegel, J-C. Pellat, et R. Rioul, *Grammaire méthodique du français*, PUF, Pais, 1994, p.110.

l'explique Picoche s'il faut traiter les unités en « *disjonction* », l'une séparée de l'autre, enregistrées sous des entrées différentes, ou s'il faut les voir plutôt comme un « *cas de subduction* », où un sens est « *dérivé* »[56] d'un autre.

Contrairement aux unités homonymiques, les unités polysémiques partagent un fond sémantique. Ainsi, par exemple nous dirons que le terme « bureau » est polysémique parce que ses différents sens, tels que le bureau meuble sur lequel nous écrivons et la pièce contenant ce type de meuble ou le lieu du travail, proviennent d'un même sens premier. Quant à l'homonymie qui caractérise des mots radicalement distincts dont la forme est la même.

Ainsi, nous considérons que le terme « avocat » qui signifie l'homme de loi et « avocat » le fruit, sont deux homonymes. Si nous cherchons l'étymologie des deux termes nous trouvons que le premier provient du latin *advocatus,* tandis que le second provient du *nahuatl* qui est une langue indienne. Selon les exemples que nous venons d'avancer nous dirons que la polysémie se définit par un critère diachronique et étymologique. Or, considérer une unité lexicale comme polysémique ou homonymique selon le critère diachronique n'est pas une tâche aisée et qui dans certains cas serait une recherche en vain.

Nous concluons par ce que nous venons d'avancer ci-dessus, que la différence entre l'homonymie et la polysémie se résume comme suit : la polysémie est une unité linguistique qui contient plusieurs signifiés, et que ces signifiés contiennent suffisamment de sèmes en commun ou qu'il soit possible de dériver l'un des sens à partir d'un autre. Quant à l'homonymie, c'est l'ensemble des unités linguistique qui partagent la même apparence graphique et phonique mais qui ne partagent aucun sème, c'est-à-dire elles ne partagent aucun lien sémantique.

Si nous avons choisi, dans notre recherche, de traiter la polysémie et non pas l'homonymie c'est parce que nous avons jugé, après de maintes lectures, que la polysémie constitue un obstacle pour le Traitement Automatique du langage Naturel par rapport à l'homonymie.

III-3-Les types de polysémie

Les linguistes et sémanticiens ont parlé de plusieurs types de polysémie, et cela en fonction de plusieurs critères. Ainsi, nous essayerons de présenter les différents types de la polysémie.

---

[56] - J. Picoche, *Précis de lexicologie française*, Ed Nathan, Paris, 1977, p.55.

III-3-1- Polysémie verticale

La polysémie verticale a été décrite par Gevaudan[57]. Selon lui, elle désigne la possibilité pour un terme de s'englober lui-même, d'être son propre hyperonyme. Pour mieux comprendre ce que Gevaudan désigne par polysémie verticale, nous citons l'exemple du mot « homme ». Ce dernier sert à désigner, d'une part, le sexe masculin qui fait partie de l'espèce (concept englobé), et d'autre part il sert à désigner l'espèce humaine dans son ensemble (concept englobant).

Nous tenons à signaler que dans le cas de la polysémie verticale, les deux acceptions ne peuvent pas se rencontrer dans un même énoncé. Par exemple le mot « homme » ne peut pas s'employer pour les deux significations dans le même énoncé : « Dans les hommes, il y a les hommes et les femmes. » mais dans le cas de l'emploi simultané des deux notions, l'une d'entre elles sera probablement remplacée par un synonyme. Ainsi, pour le même énoncé cité ci-dessus nous remplaçons le mot homme dans la première acception par son synonyme « l'espèce humaine » donc, nous dirons « Dans l'espèce humaine, il y a les hommes et les femmes. ».

III-3-2- Polysémie logique

Caractérisée par James Pustejovsky[58], la polysémie logique désigne le phénomène de cohabitation entre deux acceptions en apparence contradictoire mais liées pour le même objet linguistique. Pour mieux comprendre ce type de polysémie, nous citons l'exemple de « livre » dans les deux énoncés suivants :

- *Le livre est rouge*
- *Le livre est intéressant*

Dans la première phrase Gevaudan parle d'un *artefact*[59]. Ce dernier possède une forme, une apparence, des couleurs et toutes autres caractéristiques de son état. Quant à la deuxième phrase l'unité linguistique « livre » fait référence au contenu. Ainsi, les deux acceptions, de la même unité dite polysémique, ne peuvent pas toujours être employées conjointement dans la même occurrence.

III-3-3-Polysémie contextuelle

---

[57] - P. Gevaudan, *La polysémie verticale : Hypothèses, analyses et interprétations*, PhiN. Philologie im Netz2/a997, p.21.

[58] - J. Pustejovsky, *La contrainte aspectuelle et logique de la polysémie*, presses Clarondon, 1996, p.150.

[59]- P. Gevaudan, Op.cit. p19.

En 1997, une nouvelle conception de la polysémie est apparue. Ainsi, partant du principe qu'il n'y a pas deux situations d'énonciation équivalentes, autrement dit il n y a pas deux énoncés à sens identique, Récanati[60] affirme que la polysémie est la faculté qu'ont les contextes de ne jamais être semblables à aucun autre.

La polysémie contextuelle est, donc, définie en rapport avec le contexte. Cette même notion est reprise par Gosselin en 2001, il affirme que « *par polysémie contextuelle, on désigne le fait qu'un marqueur… puisse prendre des significations au moins partiellement différentes en fonction des contextes linguistiques dans lesquels il se trouve.* »[61]

Nous trouvons aussi beaucoup d'autres appellations données à une unité polysémique. Mais nous n'allons pas s'étaler dans chaque type. Nous pouvons citer, par exemple, la polysémie complémentaire, désordonnée, et la polysémie de raffinement.

III-4- Polysémie et traitement automatique des langues

Nous avons vu dans le chapitre précédent ce que le traitement automatique du langage naturel et nous avons vu dans ce chapitre ce qu'est la polysémie. Nous avons dit qu'une unité est polysémique lorsqu'une unité lexicale renvoie à plusieurs sens. Loin de gêner pour communiquer, la polysémie est au contraire une source de souplesse et de richesse dans les langues. Or, la prise en compte de la polysémie en Traitement Automatique du Langage Naturel a été problématique. Les chercheurs en TALN se posaient la question sur la possibilité d'associer automatiquement un sens à une unité dans un énoncé donné.

Deux démarches sont à suivre pour une automatisation de l'unité polysémique. D'abord, nous devons déterminer tous les sens possibles pour chaque mot susceptible d'être désambigüisé, autrement dit la première étape consiste à regrouper tous les sens d'une unité lexicale dans différents énoncés. Ensuite, nous devons déterminer quel sens est le bon dans tel ou tel contexte. Ces deux tâches faciles pour l'homme constituent un problème pour l'ordinateur.

Deux pistes sont suivies pour la désambiguïsation du sens en TAL. En premier lieu, nous avons ceux qui utilisent les distinctions de sens présentes dans les dictionnaires. En second lieu, nous avons ceux qui utilisent les méthodes statistiques pour repérer des patterns de cooccurrence des mots en contexte comme le cas de google traduction. Ainsi, dans les méthodes statistiques, le sens le plus probable pour une occurrence d'un mot donnée est celui qui va maximiser une certaine relation d'affinité avec le contexte de cette occurrence.

---

[60]- F. Récanati, *La polysémie contre le fixisme. Langue Française*, n 113, Mars 1997. Aux sources de la polysémie nominale, P.115.

[61]- L. Gosselin, *Le traitement de la polysémie contextuelle dans le calcul sémantique*, Intellectica, 1996/1, p94.

Lesk a crée en 1986 une méthode permettant de relier des définitions si elles ont des mots en commun. Ainsi, la désambiguïsation d'un mot en contexte se fait en choisissant pour lui et les mots qui l'entourent les définitions qui se recoupent le plus. Cette méthode posait quelques problèmes, par exemple dans le cas où il y a des définitions lapidaires. Pour améliorer cette méthode Walker pense à utiliser des codes de sujet, ces derniers sont un ensemble de primitives servant à classifier les sens d'un mot par sujet.

De nos jours, les méthodes de désambiguïsation sémantique sur corpus se développent. Cette méthode consiste à analyser les mots qui co-occurrent avec les mots polysémiques sur des corpus à grande échelle. Google traduction utilise aussi cette méthode, son système s'entraine à modéliser le sens de chaque mot, en fonction de leur contexte à partir de corpus d'exemples sémantiquement étiquetés. Ensuite la méthode statistique est utilisée pour calculer le bon sens à partir d'une mesure de similarité entre les caractéristiques des sens modélisés et celles du contexte de l'occurrence considérée.

III-5- Polysémie et traduction

Les différences qui existent entre les langues apparaissent plus nettement dès que nous mettons en correspondance des unités lexicales. Pierette Bouillon affirme dans l'ouvrage collectif *La traductque* que « *l'analyse des langues naturelles heurte en effet au problème de l'ambiguïté des langues naturelles, ambiguïté à la fois lexicale et structurale. On parle d'ambiguïté lexicale, quand un mot peut recevoir plusieurs catégories syntaxiques ou significations. Ce terme englobe aussi les ambiguïtés bilingues (ou contrastive) qui désignent que deux mots qui, tout en étant la traduction l'un de l'autre, ne dénotent pas exactement le même concept ou le même ensemble de concepts.* Par ailleurs, *le terme d'ambiguïté structurale fait référence à une phrase qui peut être dotée de plusieurs représentations.* »[62]. Pour mieux expliquer les propos de Bouillon, nous dirons que dans ce passage Bouillon parle de deux ambiguïtés : en premier lieu, nous avons les ambiguïtés du langage naturel, parce que ce dernier n'est pas stable. En second lieu, nous avons les ambiguïtés dites bilingues, c'est-à-dire que lors du passage d'une langue à une autre dans un contexte bilingue, nous somme confrontés à plusieurs ambiguïtés dues au fait que chaque langue a sa propre vision du monde.

Ainsi, la polysémie se caractérise par son état instable, nous savons que celle-ci est dynamique, et qu'elle change de sens dès que nous changeons de contexte ou cotexte. Donc, nous parlons ici d'ambiguïté des langues. En outre, la polysémie constitue aussi un problème

[62] - P. Bouillon, *La traductique*, Les presses de l'université de Montréal, AUPELF UREF, 1993, p.19.

fondamental en traduction. Chaque langue à sa vision du monde, et un terme n'a jamais le même sens d'une langue à une autre. Donc, nous parlons ici d'ambiguïté bilingue.

Par ailleurs, les chercheurs en traduction automatique ne parvenaient pas à amener un ordinateur à savoir de quoi parle telle ou telle phrase. Comme exemple du genre de problème que les ordinateurs étaient hors d'état de résoudre, citons celui qui consiste à attribuer les sens à chaque terme dans les deux phrases suivantes :

- ✓ La pièce est dans le sac
- ✓ Le sac est dans la pièce

Si nous essayons de traduire ces deux phrases sur Google traduction nous aurons les deux phrases suivantes

- ✓ جزء في كيس
- ✓ الحقيبة هو في جزء منه

Ainsi, nous remarquons que la traduction qui a été donnée par Google est mauvaise, et ceci est dû au fait que pour comprendre de telles phrases, il faut mobiliser des connaissances sur la taille relative des choses dans le monde réel : celle d'un sac pouvant contenir la pièce de la monnaie ; celle d'un sac par rapport à une pièce, c'est-à-dire une chambre, pouvant le contenir.

PARTIE PRATIQUE
LE TALN : VERS UNE IMPOSTURE DES MOTS

Préambule

Après avoir présenté, dans la partie théorique, les grands traits du traitement automatique du langage naturel et exposé les différentes théories qui définissent la polysémie. Nous allons passer à la partie pratique de ce travail, où nous allons présenter, dans un premier temps, notre corpus d'étude. Puis nous entamerons, dans un second temps, les deux chapitres qui constituent cette partie.

Ainsi, nous allons expliquer via le premier chapitre s'intitulant « l'existence du sens et/ou des sens », les différents sens d'une unité polysémique et cela par le biais de notre corpus. Nous allons expliquer, en nous basant sur la théorie connotative de Jean Michel Adam, le passage d'une unité polysémique de la signification au sens. Par ailleurs, nous passerons à l'analyse des unités polysémiques dans notre corpus.

Puis, nous entamerons, dans un second chapitre s'intitulant « D'une langue étrangère à l'étrangeté du langage », le processus de traduction automatique sur Google translate. , nous essayerons d'expliquer les différents problèmes qu'il a rencontrés lors de sa traduction des unités polysémique. Enfin, nous conclurons par l'étranger et/ou l'étrangeté dans la polysémie en TALN.

Google translate

Depuis le début du XX ème siècle, nous affirmons que nous sommes, de plus en plus, à l'ère de la technologie de l'information. Des réseaux de communications rendent le contact entre les individus plus facile et rapide. Ainsi, parmi ces réseaux, nous citons Internet. Ce dernier occupe une place importante dans la vie de tous les jours parce qu'il la rend plus facile.

Un grand moteur de recherche Google occupe le sommet de tous les réseaux sur Internet. Ainsi, Google nous permet de faire des recherches et toucher du doigt des documents autrefois inaccessibles. Par ailleurs, Google tente de remplacer l'Homme dans ses différents domaines dont le plus important est la traduction.

« Google » est un mot dérivé du terme anglais « googol » qui désigne le chiffre 1 suivi de 100 zéros. Ce terme traduit bien le but d'organiser un volume infini d'informations sur le web. Ainsi, Google nait d'un coup de génie de deux jeunes nord-américains : Larry Page et Sergey Brin.

Google translate est un outil de traduction gratuit disponible sur Internet. Il nous permet de traduire instantanément des phrases, des documents, et même des sites web entiers. Ainsi, si nous avons choisi de mener notre étude sur *Google translate* et non pas sur un autre logiciel de traduction ; c'est par ce qu'il est le moteur le plus avancé en matière de traduction automatique.

Par ailleurs, dès que nous traduisons des documents sur Google translatent, nous avons l'impression qu'une armée multilingue travaille pour Google. Or, ces traductions obtenues ne sont, en réalité que l'œuvre d'ordinateur. Ce dernier utilise un processus appelé « *traduction automatique statistique* »[63] qui n'est qu'une expression pour dire que, les ordinateurs génèrent des traductions sur la base des modèles récupérés à partir de quantité impressionnantes de textes.

Pour les créateurs de *Google translate*, un ordinateur peut apprendre une langue de la même façon qu'un être humain. Si nous volons enseigner une langue à quelqu'un, nous commencerons sans doute par lui donner le vocabulaire nécessaire et par lui apprendre les règles grammaticales grâce auxquelles il pourra construire une phrase. De même, un ordinateur peut apprendre une langue en se référant au vocabulaire et un ensemble de règles.

[63] - David Bellos*, Le poisson et le bananier : Une histoire fabuleuse de la traduction*, Ed Flammarion, 2012, p.172.

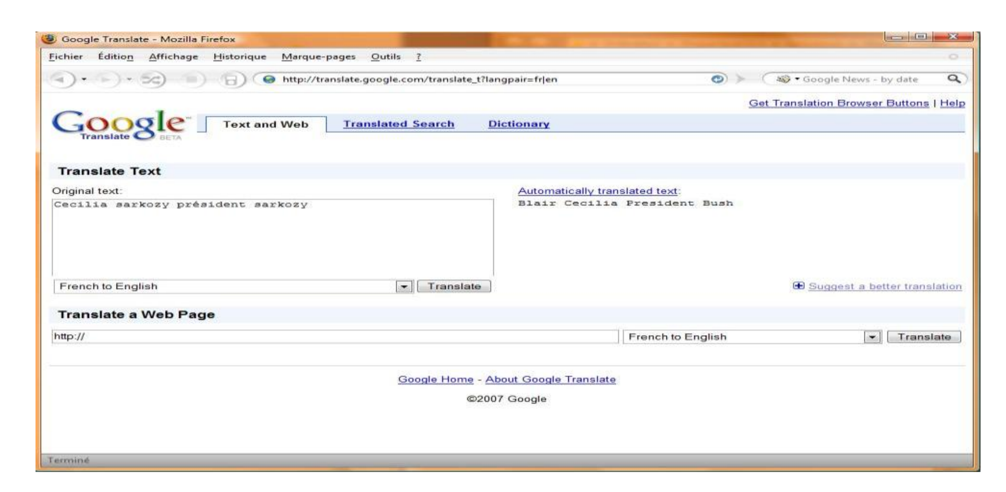


Image correspondante à Google traduction

II-L'article journalistique

L'écriture est un système de représentation de la parole et de la pensée par des signes graphiques conventionnels. Elle est devenue un système de communication humain. Ainsi, toute personne qui écrit agit pour être lue par une autre personne ou par plusieurs destinataires.

Nous avons plusieurs types d'écritures, parmi elles, nous citons l'écriture journalistique qui constitue le corpus étudié dans notre travail de recherche. Ainsi, un article journalistique est destiné à être publié dans le but d'informer le public. Par ailleurs, l'écriture journalistique est destinée à transmettre des informations étayées, vérifiées, recoupées. Dès lors, le journaliste se doit de connaitre parfaitement le sujet dont il parle avant de transmettre ses informations aux lecteurs.

Or, comme nous pouvons le constater, la place attribuée à chaque article est trop restreinte pour permettre au journaliste d'écrire son article d'une manière spontanée. Pour cela le texte journalistique se caractérise par un style connoté. Ainsi si nous avons choisi le texte journalistique, c'est pour une seule et unique raison : l'existence de mots polysémiques

Nous avons choisi pour notre étude deux articles du journal algérien *Le quotidien d'Oran*[64], ce choix du journal s'explique par le fait que nous avons trouvé des articles intéressants pour la traduction. Ainsi, Le Quotidien d'Oran est un journal algérien de langue française. Le premier numéro est apparu en 1994 et il était fondé par un groupe de citoyens.

---

[64] -http://www.lequotidien-oran.com/


LE TRAIN ALGER-TIZI OUZOU SUSPENDU POUR RAISONS SÉCURITAIRES...

Le Quotidien

30 blessés et 25 arrestations à El-Ançor

VIOLENTES ÉMEUTES PRÈS D'ORAN...

ORAN / ALICANTE / ORAN

14 000 DZD

Image correspondante au journal Le quotidien d'Oran

Les articles que nous avons choisis pour notre corpus ont été publiés le 15 octobre 2012 dans la rubrique culture. Ainsi, le premier article choisi s'intitule : « Kamel Bouchama au *Le Quotidien d'Oran* : «Si les hommes sont oublieux, l'histoire a bonne mémoire »[65], écrit par El-Houari Dilmi. Quant au deuxième article, il s'intitule : « Théâtre et histoire coloniale : l'Algérie en scène »[66] écrit par Ahmed Cheniki.

Pour se faire, nous allons traduire ces deux articles sur *Google translate*. Ensuite, nous étudierons les mots polysémiques. La traduction va se faire de la langue française vers la langue arabe standard. Par ailleurs, le choix de la langue arabe au détriment d'autres langues ne peut être justifié que par le fait que nous la maitrisons plus que d'autres dont nous ne possédons que les rudiments.

III- Langue de traduction

III-1- La langue française

La langue française est une langue romane. Sa grammaire et sa plus grande partie de son vocabulaire sont issues du latin. Au Moyen Age, la langue française est faite d'une multitude de dialectes dont nous distinguons principalement les parlers d'oïl et les parlers d'oc.

Aujourd'hui, la langue française fait partie des langues les plus importantes au monde. C'est la langue maternelle des Français et la seconde langue dans plusieurs pays tels les pays du grand Maghreb, en l'occurrence l'Algérie.

Par ailleurs, la langue française fait partie des premières langues automatisées. Elle fait partie également des langues traduites par *Google translate*.

---

[65] -http://www.lequotidien-oran.com/index.php?news=5158631 15 octobre 2012.

[66] -http://www.lequotidien-oran.com/index.php?news=5174162 15 octobre 2012.

III-2- La langue arabe standard

L’arabe, langue du Coran, est dite arabe classique. Investie d’une notion de « sacrée », elle remplit une fonction liturgique. Possédant une grammaire fixée, elle remplit une fonction communicative.

En Algérie, l’arabe est la langue nationale du pays. Autrement dit, elle est la forme linguistique utilisée dans les institutions étatiques. Cette langue a évolué au contact des langues indo-européennes, notamment aux niveaux lexical et morphosyntaxique afin de répondre aux nouvelles exigences du monde moderne. La langue arabe fait partie des langues étudiées pour la traduction automatique, notamment, par *Google traduction*.

Dans notre étude nous nous intéresserons non pas à la définition et l’historique des deux langues, mais à leur structure syntaxique et sémantique. Nous dirons que la langue française et la langue arabe partagent un univers linguistique tout à fait différent. Ainsi, les lexiques des deux langues, l’une romane et l’autre sémitique, ont des origines différentes et ne partagent aucun lien de parenté.

En outre, leur syntaxe est opposée. Par exemple dans la langue arabe nous commençons la phrase par le verbe puis passer au sujet, or en français c’est l’inverse. Enfin, nous dirons aussi que ces deux langues diffèrent même par leur écriture. Ainsi, elles possèdent des caractères différents et l’écriture arabe commence de la droite vers la gauche quant à l’écriture française elle va de la gauche vers la droite

IV- La traduction des articles sur Google traduction

IV-1- L’article n 1

Kamel Bouchama au « Le Quotidien d'Oran »: «Si les hommes sont oublieux, l'histoire a bonne mémoire»

L'ancien ministre de la Jeunesse et des Sports, Kamel Bouchama, reconverti avec un certain succès en un «historien du futur antérieur» comme il se qualifie lui-même, était jeudi à Tiaret où il a animé une conférence-débat autour de son dernier livre: «Kaïd Ahmed, un homme d'Etat». Nous l'avons approché pour lui poser quelques questions auxquelles il a bien voulu répondre sans détours ni fioritures.

Le Quotidien d'Oran: Pourquoi ce livre sur Kaïd Ahmed et pourquoi à ce moment précis de l'histoire du pays ?

Kamel Bouchama: L'écriture de ce livre procède d'abord d'une motivation première, celle de participer à l'écriture de l'histoire de mon pays. Et puis pour répondre est-ce qu'il faut réhabiliter Kaïd Ahmed ou pas, je vous dis tout de go que mon intention n'est pas de réhabiliter ce grand monsieur puisque c'est l'histoire elle-même qui l'a fait à ma place. Parce que c'est justement l'Histoire avec un grand «H» qui n'oublie pas ses hommes auxquels elle est la première à rester fidèle. Je veux dire par là si les hommes sont oublieux, l'histoire a bonne mémoire. A travers le contenu de mon livre, j'ai surtout voulu mettre en exergue le caractère exceptionnel de cet illustre personnage et ce sur tous les plans, que ce soit sur l'orientation qu'il a donnée à son action alors que le pays sortait exsangue d'une très longue nuit coloniale, ses conceptions en matière de gestion politique et économique du pays ou encore le rêve inaccompli qu'il nourrissait de voir l'Algérie, en un temps relativement court, arrimée au train à grande vitesse du développement au sens où l'entendaient les hommes faits de la même pâte que lui et même les dirigeants politiques de son époque.

Q.O.: Ne pensez-vous pas que votre livre intervient à un moment où le pays est toujours en quête d'une écriture décomplexée et surtout «dépassionnée» de son histoire ?

Bouchama كامل في "لو كوتيديان دوران": "إذا كان الرجال هم النسيان، قصة لديه ذاكرة جيدة"

وكان الوزير السابق للشباب والرياضة، كامل، وتحويلها مع بعض Bouchama النجاح في "مؤرخ أفضل في المستقبل" كما يسميه بنفسه، على تيارت الخميس حيث انه أدار حلقة نقاش حول أحدث كتاب: "قايد أحمد، رجل دولة". تناولنا لأسأله بعض الأسئلة التي أجاب بصراحة التكرم أو زخرفة.

لو كوتيديان دوران: لماذا هذا الكتاب ولماذا أحمد قايد في هذه المرحلة من تاريخ بلدنا؟

كامل: كتابة هذا الكتاب لأول Bouchama مرة ينفذ الدافع الرئيسي للمشاركة في كتابة تاريخ بلدي. والرد على ما لإعادة تأهيل أحمد قايد أم لا، وأنا أقول لكم بصراحة أن نيتي ليست لاستعادة هذا الرجل العظيم لأنه هو القصة نفسها لم في مكاني. لأنه هو رأس المال الذي لا "H" بالضبط التاريخ مع ينسى رجاله لأنها من وأول من يكون وفيا. أعني، إذا كان الناس هي النسيان، والقصة لديه ذاكرة جيدة. من خلال محتويات كتابي، أردت خاصة لتسليط الضوء على الطابع الاستثنائي هذه الشخصية اللامعة وعلى جميع المستويات، سواء كان الاتجاه الذي أعطى لعمله حيث ان البلاد خرجت نزف استعمارية ليلة طويلة جدا على التصاميم إدارته السياسية والاقتصادية للبلد أو حلم لم يتحقق أنه آوى ترى الجزائر في وقت قصير نسبيا، تأمين التنمية قطار فائق السرعة بمعنى يفهمها الرجال البالغين من نفس الاشياء عليه والقادة السياسيين حتى من وقته.

OQ: ألا تعتقد أن كتابك يأتي في الوقت الذي ما زالت البلاد بحثا عن الكتابة غير مأهولة وخاصة "نزيه" في تاريخها؟

K.B.: Justement, nous sommes à la veille de la célébration du cinquantième anniversaire de l'indépendance du pays. Et parce que le message de l'histoire est d'abord destiné aux franges les plus jeunes de notre société, j'ai écrit ce livre pour non pas aller dans le sens de ceux qui disent que les révolutions commencent par dévorer les hommes qui les pensent mais juste pour marquer une halte salutaire et dire à quel point des hommes dignes de cette Algérie ont été victimes de campagnes et autres manoeuvres très dures et toujours enfoncés encore plus profond, même enterrés dans leurs tombes.

Q.O.: Justement, pour revenir sur la question de la réhabilitation, selon vous, Kaïd Ahmed doit-il être réhabilité ou l'a-t-il déjà été au travers de votre livre ?

K.B.: Non je ne crois pas, personne n'a ou n'a pas à le faire pour réhabiliter le commandant Si Slimane. La question ne se pose pas en ces termes mais s'impose comme un devoir de mémoire qui ne peut qu'être juste vis-à-vis des artisans de l'histoire comme l'enfant digne de Tiaret. Je vous ai déjà dit que l'histoire elle-même sait réhabiliter ceux qui le méritent mais, pour moi, mon livre se veut comme une sorte de remémoration d'un grand homme qui a vécu comme un fleuve impétueux qui se tarit parce que quelqu'un a voulu casser le robinet. Pourquoi ? L'histoire est encore interpellée pour nous le faire savoir un jour qui viendra, c'est du moins là l'espoir que nourrissent les jeunes qui auront à gérer ce pays un jour qui viendra.

Q.O.: Mais pourquoi, selon vous, l'homme continue toujours de souffrir d'une sorte d'ostracisme «tacite» plus de trente-trois ans après sa disparition ?

KB: في الواقع، ونحن في عشية الذكرى السنوية الخمسين لاستقلال البلاد. ولأن المقصود في المقام الأول رسالة من القصة لصغار هامش المجتمع لدينا، كتبت هذا الكتاب لعدم ذهاب في اتجاه أولئك الذين يقولون أن الثورات تبدأ تلتهم الرجال الذين أعتقد ولكن فقط لجعل حد ويقول نافع كيف يستحق أن الرجال وقعوا ضحايا للحملات الجزائر والمناورات من الصعب جدا ودفعت دائما أعمق، ودفن حتى في قبورهم.

QO: بالضبط، للعودة إلى مسألة إعادة التأهيل، وفقا للكم، أحمد قايد أن إعادة تأهيل أو أنه قد تم بالفعل من خلال كتابك؟

KB: لا أنا لا، لا أحد فعل أو لم يفعل لإعادة تأهيل القائد سي سليمان. مسألة غير واردة في هذه الشروط لكنها تفرض واجب الذاكرة التي لا يمكن إلا مجرد وجها لوجه مع صناع التاريخ في مرحلة الطفولة تستحق تيارت. لقد سبق أن قال لك القصة نفسها يعرف تأهيل أولئك الذين يستحقون ذلك، ولكن بالنسبة لي، فإن المقصود كتابي كنوع من ذكرى رجل عظيم عاش باعتبارها نهر متهور الذي يجف بسبب شخص أراد واحد لكسر الصنبور. لماذا؟ ما تزال تواجه تحديات القصة لإعلامنا يوم سيأتي، على الأقل هناك أمل في أن إطعام الصغار الذين سيتولى إدارة هذا البلد في يوم من الأيام سوف يأتي.

OQ: لماذا تعتقد أن الرجل لا يزال يعاني من نوع من النبذ "ضمني" سنوات أكثر من 33 بعد وفاته؟

K.B.: Cela est la conséquence de la culture que nous avons dans ce pays, c'est-à-dire la culture de l'oubli. Il y a même pire, celle que l'on appelle la culture du mépris. Mais bien au-dessus de tout cela, je sais que Kaïd Ahmed de son vivant ne voulait jamais répondre par le mépris à ceux qui n'ont eu de cesse de le dénigrer juste parce que sa tête ne cadrait pas avec le «casting» de l'époque. Le commandant Slimane se plaçait toujours plus haut que ces attaques gratuites et orientées au point qu'il se prenait lui-même pour un «décalé» tellement tout lui paraissait décalé autour de lui. Et j'en veux pour preuve, ses pairs au sein du Conseil de la Révolution qui savaient tous que l'enfant de l'antique Tihert avait quelque chose de «spécial» par rapport à la bien-pensance de l'époque. Pour moi, celui qui a dit que l'Algérie est au bord du précipice mais elle a fait un grand pas en avant est tout simplement un visionnaire. Un homme en avance sur son temps et qui savait mieux que quiconque ce que voulait dire les trois révolutions de l'époque.
Q.O.: Vous avez été vous-même formé à l'école de Kaïd Ahmed, que vous inspire ce personnage trente-trois ans après sa mort ?
K.B.: J'ai appris en fréquentant son école, comme vous dites, ce que veut dire courage, volonté et surtout don de soi-même. Moi, personnellement, c'est avec lui que j'ai appris ce que voulait dire le sens de l'initiative et la nécessité de combattre cette satanée «chose» que l'on appelle l'expectative. Kaïd Ahmed était de cette race qui abhorrait de participer à une bagarre sans y prendre part d'une manière ou d'une autre.
Q.O.: Savez-vous quelque chose de précis sur les raisons réelles de sa mort ?

KB: هذا هو نتيجة للثقافة لدينا في هذا البلد، وهذا هو القول، وثقافة النسيان. بل هناك ما هو أسوأ، ثقافة ما يسمى ازدراء. لكن قبل كل شيء، وأنا أعلم قايد أحمد في حياته لم يكن يريد أن تستجيب بازدراء أولئك الذين كانوا باستمرار لمجرد تشويه سمعة رأسه لم تناسب "الصب" من الوقت. قائد سليمان وضعت دائما فوق هذه الهجمات لا مبرر له والتوجه إلى النقطة التي تولى لنفسه "متداخلة" بحيث بدا تحول كل شيء من حوله. وأريد أن أثبت أقرانه في مجلس الثورة يعرف كل أن الطفل من Tihert القديمة كان شيئا "خاصة" بالمقارنة مع البر الذاتي من الوقت. بالنسبة لي، من قال أن الجزائر على وشك ولكنها اتخذت خطوة كبيرة إلى الأمام هو مجرد رؤية. عرف رجل قبل وقته وأفضل من أي شخص ما يعني ثلاث ثورات في ذلك الوقت.

QO: كنتم نفسك تلقى تعليمه في مدرسة أحمد قايد، كنت هذه الشخصية مستوحاة ثلاثة وثلاثين عاما بعد وفاته؟

KB: تعلمت من الذهاب إلى المدرسة، كما تقول، وهذا يعني الشجاعة والعزيمة وقبل كل شيء هبة الذات. لي، شخصيا، فمن معه على أن تعلمت ماذا يعني أن روح المبادرة والحاجة إلى محاربة توقع 'الشيء' لعنة تسمى. وكان قايد أحمد هذا السباق الذي يمقت المشاركة في معركة من دون أخذ جزء في شكل او في آخر.
OQ: هل تعرف أي شيء محدد حول الأسباب الحقيقية لموته؟

K.B.: Vous savez, je ne fais pas partie de ceux qui cherchent des poux dans la tête des autres. Mais je dois vous dire que les circonstances de sa mort sont relatées dans mon ouvrage sur la base de témoignages vérifiés et vérifiables. Pour être franc avec vous, Kaïd Ahmed est mort de son exil injuste. Il est parti parce qu'il a trop aimé sa patrie (sans jeu de mots) jusqu'à la mort, même si personne aujourd'hui n'est capable de dire avec certitude de quoi est mort cet homme si particulier.

K.B.: Vous savez, je ne fais pas partie de ceux qui cherchent des poux dans la tête des autres. Mais je dois vous dire que les circonstances de sa mort sont relatées dans mon ouvrage sur la base de témoignages vérifiés et vérifiables. Pour être franc avec vous, Kaïd Ahmed est mort de son exil injuste. Il est parti parce qu'il a trop aimé sa patrie (sans jeu de mots) jusqu'à la mort, même si personne aujourd'hui n'est capable de dire avec certitude de quoi est mort cet homme si particulier.

Q.O.: En tant qu'ancien commis de l'Etat, quelle est votre opinion sur les réformes politiques engagées dans le pays et qu'en attendez-vous ?

K.B.: Personnellement, j'ai une opinion tranchée sur le sujet. Pour moi, il ne s'agit pas de faire changer le costume lorsqu'il devient trop ample ou trop étroit. Vous n'avez qu'à observer ce qui se passe actuellement dans les pays arabes et le prix à payer pour une orientation résolue dans le sens de la marche de l'histoire moderne universelle. Les révolutions dans les pays voisins et autres ne sont pas des parties de plaisir au sens que la notion fallacieuse de réformes pour réformer ne veut plus rien dire, y compris pour ceux qui réfléchissent à trois fois avant d'ouvrir leur bouche.

KB: أنت تعرف، أنا لست واحدا من أولئك الذين يسعون القمل في أذهان الآخرين. ولكن لا بد لي من ان اقول لكم ان وروى ظروف وفاته في كتابي على أساس الحسابات المدققة ويمكن التحقق منها. أن نكون صادقين معكم، توفي أحمد قايد غير عادلة في المنفى. غادر لأنه يحب بلده جدا (لا يقصد التوريه) حتى الموت، حتى لو لم يكن أحد اليوم يستطيع أن يقول على وجه اليقين ما هو خاص جدا حول هذا الرجل مات.

KB: أنت تعرف، أنا لست واحدا من أولئك الذين يسعون القمل في أذهان الآخرين. ولكن لا بد لي من ان اقول لكم ان وروى ظروف وفاته في كتابي على أساس الحسابات المدققة ويمكن التحقق منها. أن نكون صادقين معكم، توفي أحمد قايد غير عادلة في المنفى. غادر لأنه يحب بلده جدا (لا يقصد التوريه) حتى الموت، حتى لو لم يكن أحد اليوم يستطيع أن يقول على وجه اليقين ما هو خاص جدا حول هذا الرجل مات.

QO: ككاتبة السابق للدولة، ما هو رأيك في إصلاحات سياسية في البلاد، وماذا تتوقع؟

KB: شخصيا، لدي رأي قوي حول هذا الموضوع. بالنسبة لي، ليس لتغيير الزي عندما يصبح واسعة جدا أو ضيقة جدا. أنك لا تملك إلا أن ننظر إلى ما يحدث في الدول العربية وثمن التوجه حلها في اتجاه السفر من التاريخ الحديث للجميع. الثورات في الدول المجاورة وأجزاء أخرى ليست متعة، بمعنى أن المفهوم الخاطئ للإصلاح من الإصلاحات لا يعني شيئا بعد الآن، بما في ذلك أولئك الذين يعتقدون ثلاث مرات قبل فتح أفواههم.

J'ai la conviction ferme qu'à trop vouloir replâtrer la façade, on finit par laisser apparaître la grosse montagne de poussière cachée sous le tapis. En un mot, et à mon humble avis, les Algériens ont besoin de voir des idées nouvelles incarnées par des visages nouveaux intervenir dans la gestion de leur destin commun, à moins de huit mois du 50e anniversaire du retour du soleil de la liberté sur un pays qui a de tout temps suscité la jalousie maladive des uns et la félonie impardonnable des autres.

لدي قناعة راسخة، والرغبة الجص أكثر من واجهة، ونحن في النهاية تكشف عن الجبل الكبير من التراب مخبأة تحت السجادة. في كلمة واحدة، وفي رأيي المتواضع، فإن الجزائريين بحاجة إلى أن نرى أفكار جديدة يجسدها الوجوه الجديدة المشاركة في إدارة مصيره المشترك في غضون ثمانية أشهر من الذكرى 50 لعودة الشمس من الحرية على وقد اجتذب هذا البلد دائما الغيرة المرضية من جناية لا تغتفر كل الأخرى.

Article n 02

Théâtre et Histoire coloniale: l'Algérie en scène

Le théâtre, comme d'ailleurs les autres formes de représentation littéraire et artistique, a souvent accompagné les grandes expériences révolutionnaires.

La Révolution de 1789, la Commune de Paris, la Révolution russe, la guerre civile espagnole et les luttes de libération en Afrique et en Amérique Latine ont été soutenues par des artistes et des écrivains qui ont cherché à donner à voir le vécu des patriotes et des républicains s'opposant à divers espaces répressifs et dictatoriaux. Ainsi, l'Histoire côtoie les aventures artistiques et les investit d'un contenu particulier et leur permet de mettre en œuvre de nouvelles configurations esthétiques. Brecht, Adamov, Meyerhold, Vitez, Mnouchkine ou Kateb Yacine ont parcouru des pans entiers de l'Histoire et ont tenté de montrer que les jeux artistiques, contrairement aux prétentions structuralistes, ne peuvent, en aucune manière, se désolidariser des jeux et des enjeux idéologiques et politiques.

L'Histoire et l'art sont au cœur de la mêlée politique, artistique et esthétique. Kateb Yacine propose une lecture fondamentalement politique du combat anticolonial dans Le cadavre encerclé, convoquant une manière de faire faisant cohabiter tragique et épique, mythe et Histoire, à l'instar de Césaire. Il se trouve, et c'est logique, que ces deux auteurs ont profondément séduit l'homme providentiel, Jean-Marie Serreau, qui a fait connaitre en France, entre autres écrivains de théâtre, Bertolt Brecht et de jeunes dramaturges africains qu'il a mis en scène, malgré de nombreuses menaces et les dures récriminations des milieux colonialistes qui n'admettaient pas que le théâtre témoigne de la tragédie des peuples africains, tout en prenant position pour les victimes de la sauvagerie coloniale dont les chantres cherchent aujourd'hui à se refaire une virginité et à nier ce moment sale de l'Histoire «occidentale».

مسرح التاريخ الاستعماري والجزائر في المشهد

المسرح، وأشكال أخرى من التمثيل في الواقع الأدبي والفني، كثيرا ما يكون مصحوبا التجارب الثورية الكبيرة.

وقد تم دعم الثورة عام 1789، كومونة باريس، الثورة الروسية، الحرب الأهلية الإسبانية وتحرير تناضل في أفريقيا وأمريكا اللاتينية من قبل الفنانين والكتاب الذين سعوا إلى إبراز تجارب الوطنيين و الجمهوريون معارضة مساحات مختلفة القمعية والدكتاتورية. وهكذا، تاريخ التدليك مغامرات الفنية وتستثمر على محتوى معين، وتتيح لهم تنفيذ تكوينات جمالية جديدة. وقد سافر بريشت، ادموف، Meyerhold، فيتيز، أو كاتب ياسين منوشكين مساحات شاسعة من التاريخ وحاول أن يظهر أن ألعاب الفني، على عكس المطالبات البنيوية لا يمكن، بأي شكل من الأشكال، وقطع الألعاب والتحديات الأيديولوجية والسياسية.

التاريخ والفن في قلب المعمعة السياسية للوالفنية والجمالية. كاتب ياسين يقدم قراءة سياسية في الأساس للنضال المناهضة للاستعمار في لجثة تحيط بها واستدعاء وسيلة للعيش معا أسطورة، وملحمة المأساة والتاريخ، مثل سيزير. هو عليه، وأنه أمر منطقي أن هذه الكاتبين قد فاز بعمق الرجل محظوظ، جان ماري Serreau، الذين معروفة في فرنسا، من بين الكتاب المسرحيين أخرى، برتولت بريشت والكتاب المسرحيين الشباب التي الأفريقي نظمت، على الرغم من العديد من التهديدات والاتهامات القاسية البيئات المستعمرين لم نعترف بأن المسرح يعكس مأساة الشعوب الأفريقية، بينما اتخاذ موقف لضحايا الوحشية الاستعمارية التي المطربين يسعون الآن لإعادة بناء عذريتهن، وكذلك حرمان هذه "اللحظة القذرة التاريخ "الغربية".

En Algérie, la colonisation française, succédant aux sombres moments «ottomans», va transformer radicalement les réalités quotidiennes, en neutralisant toute expression culturelle autochtone et en imposant ses propres formes, sans pour autant permettre aux Algériens d'aller à l'école, théoriquement gratuite et obligatoire depuis la loi Jules Ferry, un grand apôtre de la colonisation. En 1962, l'Algérie comptait moins de 10% de lettrés. La colonisation est un mal intégral. Les Turcs, comme les Français, ont ignoré les populations locales qu'ils ont continuellement appauvries, socialement et culturellement, provoquant de graves schismes et engendrant des ruptures et des césures préjudiciables au fonctionnement mental des Algériens.
Le théâtre, forme européenne réadaptée à partir de l'héritage grec, comme les autres formes de représentation, va investir certains territoires urbains et séduire quelques personnes qui n'hésitèrent pas à lui fournir un contenu local tout en l'imprimant des structures artistiques locales, à tel point que nous sommes en présence de textes syncrétiques paradoxaux, juxtaposant les espaces autochtones et les jeux théâtraux, même si, bien entendu, le primat de l'appareil théâtral est évident. Ce n'est qu'au début du vingtième siècle que les Algériens, colonisés et méprisés, allaient découvrir cet art. Les premiers textes joués, à partir de 1908-1909, étaient marqués par une sorte de repli identitaire, donnant à voir un passé arabe présenté comme glorieux et des personnages historiques favorablement dépeints.
On convoque l'Histoire pour répondre au discours colonial décrivant les Algériens comme incultes et sans profondeur historique et identitaire. Ainsi, les premières pièces historiques s'inscrivaient dans ce nécessaire dialogue avec l'Autre, certes physiquement absent de l'espace dramatique, mais fortement présent, à travers cette structure dialogique, installant face à face deux figures, l'une, présente, autochtone, porteuse de valeurs historiques valorisantes et l'Autre, absente, subissant symboliquement les attaques acerbes des personnages arabes et musulmans. Ces textes, en arabe littéraire, sont certes suggestifs et peu clairs, mais proposent une autre lecture de l'Histoire et une autre conception de l'altérité.

في الجزائر، والاستعمار الفرنسي خلفا لأحلك اللحظات " العثماني" ستحول جذريا واقع الحياة اليومية، وتحييد أي تعبير السكان الأصليين الثقافية وفرض أشكال خاصة بها، دون السماح الجزائريين للذهاب إلى المدرسة، نظريا مجاني وإلزامي من القانون فيري جول، وهو الرسول العظيم من الاستعمار. في عام 1962، كان الجزائر يعرفون القراءة والكتابة أقل من 10٪. الاستعمار هو الشر الكامل. الأتراك، مثل الفرنسية، وتجاهل السكان المحليين كانوا فقراء باستمرار واجتماعيا وثقافيا، مما تسبب في الانشقاقات الخطيرة والتسبب في وقوع الواصلة يضر الجزائريين الأداء العقلي.

المسرح، وإعادة تأهيل الأوروبية من التراث اليوناني، مثل غيرها من أشكال التمثيل، وسوف تستثمر بعض المناطق الحضرية وجذب بعض الناس الذين لم يترددوا في تقديم المحتوى المحلي أثناء الطباعة في الهياكل المحلية الفني لدرجة أننا في وجود نصوص توفيقية تجاور متناقض مساحات والألعاب الأصلية المسرح، على الرغم من وبطبيعة الحال، أولوية الجهاز المسرحي هو واضح. أنها ليست سوى في بداية القرن العشرين أن الجزائريين، والاحتقار والمستعمر، واكتشاف هذا الفن. تميزت النصوص الأولى لعبت، من 1908-1909، من خلال نوع من الانعزالية، وإعطاء رؤية عربية الماضية قدمت كأحرف المجيدة والتاريخية صورت بشكل إيجابي.
واستدعى التاريخ لتلبية الخطاب الاستعماري يصف الجزائريين وغير المتعلمين ودون عمق الهوية والتاريخية. وهكذا، كانت القطع التاريخية الأولى جزءا من هذا الحوار الضروري مع الآخر، على الرغم من غياب جسديا من الفضاء مثيرة، لكن حاضرة بقوة من خلال هذه البنية الحوارية، وتركيب اثنين من الشخصيات التي تواجه بعضها البعض، واحد، ويعرض مكافأة الناقل السكان الأصليين من القيم التاريخية و، تغيب أخرى، تمر رمزيا هجمات قاسية على الأحرف العربية ومسلم. هذه النصوص والعربية الأدبية، هي بالتأكيد موحية وغير واضحة، ولكنها تشير قراءة مختلفة للتاريخ ومفهوما مختلفا للالغيرية.

Ce retournement est logique, compte tenu du contexte tragique de l'époque. Mais à partir de 1926, un glissement discursif s'opère et une rupture singulière avec les pratiques du passé va avoir lieu, mettant en œuvre la construction d'un autre discours qui demeure dual, installant deux entités antagoniques, celles du colonisé et du colonisateur. Des auteurs, parfois proches de partis nationalistes ou réformistes vont tenter de porter l'Histoire au théâtre et de décrire des situations d'exploitation tout en insistant sur la dimension nationale, mais sans aller ouvertement à la confrontation, proposant un discours voilé, travaillé par de nombreuses contorsions lexicales et des glissements sémantiques singuliers, évitant les dénonciations directes et franches. Mais tout cela allait provoquer l'ire des autorités coloniales qui réagirent négativement en censurant les pièces et les auteurs.

Cette nouvelle attitude est l'expression des profonds changements opérés dans la société algérienne qui n'admettait plus la colonisation et qui commençait à nourrir un sentiment de rejet profond de tous les symboles de la France et de l'Europe. Le théâtre ne pouvait rester en dehors de ce désir de changer les choses et de mettre en scène des situations conflictuelles, convoquant, par à-coups, des moments interstitiels, donnant à voir deux formations discursives radicalement différentes, incompatibles, dont le rapprochement est désormais impossible. Ce jeu avec l'Histoire obéissait donc à la nécessité de fournir un nouveau discours débarrassé des oripeaux de la négativité et de la conscience malheureuse.

هذا الانعكاس هو منطقي، بالنظر إلى السياق المأساوي في ذلك الوقت. ولكن من عام 1926، تحول الخطابية ويحدث كسر المفرد مع الممارسات السابقة ستعقد وتنفيذ بناء آخر لا يزال الخطاب المزدوج تثبيت كيانين العدائية، وتلك من المستعمر المستعمر ولل . عملت من الكتاب، في بعض الأحيان إلى إغلاق أو الأحزاب القومية الإصلاحية سوف محاولة لتحقيق الدراما والتاريخ لوصف حالات الاستغلال مع التأكيد على البعد الوطني، ولكن ليس بشكل علني المواجهة، وتقديم خطاب المحجبات، وذلك التشويهات العديد من التحولات المعجمية والدلالية المفرد، وتجنب الانسحاب المباشر والعلني. ولكن كل هذا يثير حفيظة السلطات الاستعمارية الذين كان رد فعل سلبيا الرقابة أجزاء من الكتاب و.

هذا الموقف الجديد هو تعبير عن التغيرات العميقة التي في المجتمع الجزائري، الذي لم يعد الاستعمار واعترف بدأت تغذي الشعور العميق رفض جميع الرموز من فرنسا وأوروبا. يمكن للمسرح أن يبقى بعيدا عن الرغبة في تغيير الأمور ومسرح الصراع، داعيا، في الزيادات، لحظات فراغي، وإعطاء التشكيلات الخطابية رؤية مختلفين جذريا، غير متوافقة، والذي هو الآن في تقريب مستحيل. يطاع هذا التاريخ وهكذا لعبة الحاجة إلى خطاب جديد خال من بهرج مع سلبية وسيئة الضمير

L'Histoire est drapée d'un ensemble lexical où deux entités antagoniques ne pouvaient se départager que par la violence et le rejet de l'Autre. Ainsi, quelques auteurs algériens vont convoquer l'Histoire pour mettre en accusation l'occupant qui détient tous les espaces de pouvoir. Les colonisés sont condamnés à occuper des lieux dépourvus de pouvoir et d'autorité. Le travail de reconquête ne peut donc se faire que par la contestation et la subversion du discours de l'Autre, trop marqué par les jeux d'autorité et de pouvoir absolu. Même les premiers hommes de théâtre algériens vont s'atteler à lorgner du côté d'une mémoire constitutive d'une Histoire autre, différente à exposer au regard du colonisateur qui n'arrête pas de fabriquer une image à sa convenance du colonisé, en se servant des savoirs d'anthropologues et d'ethnologues serviles, prêts à toute entreprise de légitimation, contribuant à la péjoration de l'Histoire des populations autochtones.
Toute la période qui a précédé la guerre de libération a consisté, chez certains auteurs «indigènes» à donner à voir une autre lecture du processus historique, démentant ainsi les jugements et les propos trop peu sérieux d'un grand nombre d'historiens et d'ethnologues attachés à la colonisation. Une relecture sérieuse des savoirs «occidentaux» et leur interprétation-interrogation est nécessaire pour mieux approcher le fonctionnement de notre société et de tous les pays anciennement colonisés. Certains auteurs algériens, certes bénéficiant d'un cursus scolaire limité, ont compris empiriquement cette réalité en tentant une inversion des structures discursives et lexicales les investissant d'un autre contenu et de nouveaux réseaux sémantiques.

ورايات التاريخ مع مجموعة المعجمية حيث كيانين عدائية لا يمكن أن تقرر من العنف ورفض الآخر. وهكذا، فإن بعض الكتاب إلى التاريخ الجزائري عقد عزل المحتل الذي يحمل جميع المساحات السلطة. استعمرت محكوم لاحتلال الأماكن دون القوة والسلطة. لا يمكن عمل ينبغي القيام به من قبل الفتح التحدي وتخريب الخطاب للآخر، شهد أيضا من لعبة السلطة والسلطة المطلقة. وحتى يتم تسخير الرجال الأولى من المسرح الجزائري إلى جانب غمز ذاكرة التاريخ يشكل واحدة مختلفة لفضح تحت المستعمر الذين لا يستطيعون التوقف عن صورة من الراحة مستعمرة من قبل باستخدام المعرفة من علماء الأنثروبولوجيا وethnologists ذليل، وعلى استعداد لأي شرعية الأعمال، والمساهمة في تدهور تاريخ الشعوب الأصلية.

طوال الفترة التي سبقت حرب التحرير كانت، في ' الأم' بعض الكتاب لإعطاء لمعرفة آخر قراءة لعملية تاريخية، وأيضا إنكار الأحكام خطيرة قليلا عن عدد كبير من المؤرخين و ethnologists تعلق على الاستعمار. وهناك حاجة إلى إعادة المعارف خطيرة " الغربية" وتفسيرها الفحص في تحسين عمل تقريبي لمجتمعنا وجميع المستعمرات السابقة. بعض الكتاب الجزائريين تستفيد بالتأكيد من منهج يقتصر فهم هذا الواقع تجريبيا من خلال محاولة انقلاب للهياكل الخطابية والمعجمية الاستثمار محتوى جديد آخر والشبكات الدلالية.

Le sens prend un glissement extraordinaire en épousant les contours d'un autre schéma politique et idéologique. La suite théâtrale de Kateb Yacine, Le cercle des représailles, constitué de deux tragédies («Le cadavre encerclé» et «Les ancêtres redoublent de férocité»), une satire (La poudre d'intelligence) et un poème dramatique est significative de cette métamorphose au niveau de la structure et du sens, permettant à la fois le dévoilement du discours colonial et la mise au jour des pratiques historiques autochtones. Le jeu avec l'Histoire se double ici de la présence des espaces mythiques et mythologiques qui transforment radicalement le discours colonial, le subvertissant et le plaçant dans une posture péjorative. Ce retournement sémantique, morphologique, lexical et syntaxique participe d'une volonté de permettre au colonisé d'être partie prenante d'une Histoire et d'un temps, certes marqué par les pesanteurs du passé, mais foncièrement ouvert. Ce n'est d'ailleurs pas pour rien que les textes «tragiques» fonctionnent comme des tragédies optimistes où la parole au singulier se nourrit d'une mémoire collective et d'un imaginaire populaire, prompts à figurer un autre discours, libéré des geôles de l'indéfini et du mythe conçu comme espace atemporel.

الاتجاه يأخذ زلة الخاصة ملامح مخطط آخر السياسية والأيديولوجية. كاتب ياسين بعد المسرح، ودائرة الانتقام، تتكون من اثنين من المآسي (" إن جثة تحيط" و " الأجداد شراسة مضاعفة")، وهي هجاء (الاستخبارات مسحوق) وقصيدة درامية خير دليل على هذا التحول في هيكل ومعنى لكشف كل من الخطاب الاستعماري وكشف الممارسات التاريخية الأصلية. اللعب مع التاريخ الزوجي عن وجود مساحات الأسطورية والأسطورية تحويل جذري الخطاب الاستعماري، وتخريب ووضعه في موقف انتقاص. هذا الانعكاس الدلالي، الصرفية، والنحوية المعجمية تنطوي على استعداد للسماح للاستعمار أن تكون جزءا من التاريخ والوقت، بالتأكيد علامة من أعباء الماضي، ولكنها مفتوحة بشكل أساسي. ليس من أجل لا شيء أن النصوص "المأساوي" وظيفة والمآسي متفائل حيث الكلمة في صيغة المفرد لتغذية الذاكرة الجماعية والخيال الشعبي، وتشمل السريع لخطاب آخر، وخالية من السجون وإلى أجل غير مسمى وأسطورة تصور كمساحة الخالدة.

Chapitre I

L’existence du sens et/ou des sens

I-Polysémie : passage de la signification au sens

Nous avons vu dans le deuxième chapitre de la partie théorique, la distinction entre le sens et la signification. Nous avons qualifié la signification de ce qui est dénoté ou explicite dans une unité lexicale ; et le sens comme tout signifié qui s'ajoute à cette signification, c'est-à-dire toutes les connotations de cette unité lexicale.

Par ailleurs, nous avons aussi défini la notion de polysémie selon différents théoriciens. Dans notre étude sur la polysémie, nous avons choisi la théorie connotative de Jean Michel Adam[67]. Dans son ouvrage, *linguistique et discours littéraire : théorie et pratique des textes*, Jeam Michel Adam définit l'unité polysémique comme la coexistence dans une même unité d'un sens dénoté et de sens connotés.

La notion de polysémie a toujours été définie avec ces deux concepts de dénotation et de connotations. Nous allons essayer d'expliquer, à partir de notre corpus, comment une unité polysémique change de sens et passe de la signification au sens et ceci par le biais de la théorie connotative de Jean Michel Adam.

I-1- Dénotation et connotation

Selon Jean Michel Adam, la dénotation est « *assimilable à la fonction référentielle du langage, elle correspond à la possibilité pour le signe de renvoyer à une réalité extra-linguistique ; elle est informative* »[68]. En revanche, les connotations sont « *considérées comme des significations supplémentaires qui se superposent à la langue décrite dans les dictionnaires et les grammaires, donnant au message une coloration spécifique* »[69].

Ainsi dans le premier article, nous avons le terme *histoire* dont nous pouvons extraire un sens dénoté et des sens connotés à ce premier sens. Dans un premier sens, le terme *histoire* signifie : réciter des évènements relatifs aux peuples ou à l'humanité en général. Par ailleurs, d'autres sens viennent s'ajouter à cette signification pour que le même terme *histoire* signifie :

- ✓ La suite de ces évènements ;
- ✓ Récit d'aventures réelles ou imaginaires ;
- ✓ Récit d'un évènement individuel.

Or la difficulté d'identifier le sens dénoté par rapport aux sens connotés n'est pas une tâche aisée. Ainsi, nous devons faire une recherche diachronique pour trouver le premier sens ou le sens dénoté. En revanche, dans certains cas, nous rencontrons des unités polysémiques dont une recherche du premier sens serait une recherche en vain.

---

[67] - Jean-Michel Adam, Op.cit. pp. 82-107.
[68] - Ibid, p.84.
[69]- Ibidem.

Nous parlons toujours d'un sens dénoté au singulier, et de sens connotés au pluriel, et ceci nous fait penser qu'il existe un sen premier, vrai autour duquel se rattachent d'autres sens figurés.

I-2- Détermination du sens polysémique

Le langage humain se caractérise par son caractère dynamique dû au mouvement perpétuel des unités lexicales, en l'occurrence les unités polysémiques. Ainsi, pour toute tentative de désambiguïsation de sens polysémique en linguistique, nous devons prendre en considération deux paramètres essentiels : le contexte et le cotexte.

I-2-1- Polysémie et contexte

Nous avons vu dans le deuxième chapitre de la partie théorique que chaque unité linguistique se rapporte à un référent dans le monde extralinguistique. Ainsi, la langue fonctionne en un étroit rapport avec le domaine extralinguistique. Par ailleurs, selon Jean Michel Adam, les connotations de l'unité polysémique *« n'existent pas en elles-mêmes : un dictionnaire de connotation est donc impensable.* »[70], nous concluons par ce passage, que les sens des unités diffèrent selon le contexte où elles apparaissent. En revanche nous ne pouvons pas regrouper toutes les connotations possibles d'une unité.

Or, pour déterminer le sens de l'unité polysémique, nous avons toujours recours au contexte. Ainsi, dans ce passage : « *En tant qu'ancien commis de l'Etat, quelle est votre opinion sur les réformes politiques engagées dans le pays et qu'en attendez-vous ?* » l'unité « *ancien* » est considérée polysémique parce qu'elle connote plusieurs sens dont une relation d'un sens dénoté est claire. Ainsi, le contexte où ce passage s'inscrit nous permet de dégager le sens exact de cette unité dans ce contexte. Par ailleurs, nous savons qu'*ancien* peut signifier :

- ✓ Qui existe depuis longtemps ;
- ✓ Qui a existé à une époque révolue ;
- ✓ Celui qui a été avant un autre dans un corps.

Ainsi, en lisant le passage, nous nous plaçons dans le contexte où il a été dit, c'est-à-dire qu'il s'agit de grade et d'une certaine hiérarchie dans un corps. De ce fait, nous pouvons déduire directement qu'il s'agit du troisième sens, c'est-à-dire, il désigne que la personne occupait avant un grade quelconque qu'elle n'occupe plus maintenant.

Par ailleurs, les sens connotés font toujours appel aux codes culturels et idéologiques. Ainsi, dans son ouvrage *Éléments de sémiologies et de mythologies,* Roland Barthes affirme

[70] - Ibid ibidem, p.84.

que *« le signifié de connotation … a un caractère à la fois général, global et diffus : c'est, si l'on veut, un fragment d'idéologie… ces signifiés communiquent étroitement avec la culture, le savoir, l'histoire, c'est par eux… que le monde pénètre le système ; l'idéologie serait en somme des signifiés de connotation.* »[71].

En outre, certains linguistes lient le langage de connotation avec la stylistique ; en l'opposant au langage de dénotation propre au langage scientifique. Autrement dit, la connotation se range du côté du langage affectif et de l'art, tandis que la dénotation se trouve dans le collectif.

I-2-2- Polysémie et cotexte

Si nous reprenons les affirmations du père fondateur de la linguistique Ferdinand De Saussure, nous dirons que chaque signe n'a de *valeur*[72] que par ses relations avec les autres signes du système linguistique. Par ailleurs, le sens des unités linguistiques est en étroit rapport avec le système de la langue qui le définit. Nous appelons cette détermination du sens des unités par rapport à ce système : le cotexte.

Jean Michel Adam affirme que la connotation *« ... est comme le produit d'une structuration textuelle du langage naturel. Le texte ne connote pas en soi, mais c'est le texte qui fait le sens connotatif du signe dans le texte »*[73].

Ainsi, outre le facteur contexte déterminant le sens de l'unité polysémique, le cotexte joue un rôle important dans l'identification des sens connotés. Par ailleurs, outre la prise en compte des facteurs extralinguistiques, la détermination du sens doit prendre en considération l'organisation interne du discours où les unités polysémiques prennent leur valeur.

Si nous voulons soulever le sens de l'unité polysémique dans l'expression « l'*histoire* coloniale de l'Algérie » nous devons faire référence aux autres unités. S'il s'agit de réciter des évènements relatifs aux peuples, de la suite de ces évènements, d'un récit d'aventures réelles ou imaginaires, ou enfin d'un récit d'un évènement individuel ; le sens se définit bel et bien par rapport aux unités « coloniale » et « Algérie ».

Ainsi, si nous demandons à un être humain de nous dégager le sens de l'unité polysémique « histoire » dans le passage cité ci-dessus, il choisira, sans doute, la suite des évènements récités sur un peuple. Ainsi, ce choix s'explique dans le système même de

---

[71]- Roland, Barthes*," Éléments de sémiologie", Communications*, 4, Ed Seuil, Paris, 1964, p.131.
[72] - Ferdinand De Saussure, Op.cit, p.181.
[73] - Jean Michel Adam, Op.cit, p. 89.

l'expression, parce que l'adjectif *colonial* et le nom du pays « Algérie » réfèrent à ce sens et non pas aux autres sens.

Nous tenons à signaler que pour la détermination du sens polysémique, c'est-à-dire s'il s'agit d'un sens dénoté ou d'un sens connoté, nous devons prendre en considération et le contexte et le cotexte. Ces derniers jouent un rôle complémentaire dans la construction du sens polysémique. Ainsi, dégager le sens de l'unité polysémique *histoire* dans le même passage cité ci-dessus « l'*histoire* coloniale de l'Algérie » nous prenons compte les deux facteurs : contexte et cotexte.

De ce fait, les unités qui entourent cette unité nous révèlent son sens ; mais aussi nous avons besoin de connaissances extralinguistiques pour déterminer le sens exact ; c'est-à-dire nous devons placer ces mêmes unités dans un contexte existant dans la réalité. Nous concluons que nous ne pouvons pas faire une étude du sens polysémique en prenant en considération juste le contexte et ne nous pouvons non plus étudier le sens en dehors du cotexte.

En revanche, Mel'cuk[74] considère que les études du sens d'une unité linguistique doivent être faites indépendamment du contexte et du cotexte. Ainsi, les études de Pottier propose de « *découvrir le sens par le repérage d'unités ultimes de signification contenues dans le mot : les sèmes* »[75], le contexte est complètement ignoré. De ce fait, le sens apparait comme un bloc homogène et indépendant.

Cette position de Mel'cuk et Pottier a été complètement remise en question, parce que nous ne pouvons pas négliger le rôle important du cotexte et du contexte dans la construction du sens dans l'unité polysémique.

Une seconde position stipule que les unités polysémiques possèdent une signification première qui se modifie sous l'action du contexte et du cotexte. Ses auteurs considèrent que le contexte et le cotexte comme un filtre qui sélectionne à travers un nombre limité de significations. Or, ce point de vue considère le sens pragmatique comme un sens additionnel ; donc il cherche à établir l'existence d'une invariante sémantique.

II-Recueille des sens polysémique dans le corpus

Nous allons présenter dans ce qui suit les unités polysémiques mal traduites par Google traduction sous forme d'un tableau ; et nous expliquerons, par la suite, les raisons pour lesquelles Google a échoué dans son opération de traduction.

---

[74] - Mel'cuk, *Vers une linguistique Sens-Texte*, Leçon inaugurale au collège de France, Chaire internationale, Les presses de l'Université de Montréal, 1997, p.10.

[75] - B. Pottier, *Linguistique Générale ; théorie et description*, Paris, Klincksieck, 1974, p.35.

II-1- Article I

L'unité polysémique en langue Source	L'unité donné par Google Traduction	La bonne traduction
l'histoire a bonne mémoire	قصة لديه ذاكرة جيدة	التاريخ لديه ذاكرة جيدة
L'histoire est encore interpellée pour nous le faire savoir un jour qui viendra, c'est du moins là l'espoir que nourrissent les jeunes qui auront à gérer ce pays un jour qui viendra.	ما تزال تواجه تحديات القصة لإعلامنا يوم سيأتي، على الأقل هناك أمل في أن إطعام الصغار الذين سيتولى إدارة هذا البلد في يوم من الأيام سوف يأتي.	ما تزال تواجه تحديات التاريخ لإعلامنا يوم سيأتي، على الأقل هناك أمل في أن إطعام الصغار الذين سيتولى إدارة هذا البلد في يوم من الأيام سوف يأتي.
Mais bien au-dessus de tout cela, je sais que Kaïd Ahmed de son vivant ne voulait jamais répondre par le mépris à ceux qui n'ont eu de cesse de le dénigrer juste parce que sa tête ne cadrait pas avec le «casting» de l'époque.	لكن قبل كل شيء، وأنا أعلم قايد أحمد في حياته لم يكن يريد أن تستجيب بازدراء أولئك الذين كانوا باستمرار لمجرد تشويه سمعة رأسه لم تناسب "الصب" من الوقت.	لكن قبل كل شيء، وأنا أعلم قايد أحمد في حياته لم يكن يريد أن تستجيب بازدراء أولئك الذين كانوا باستمرار لمجرد تشويه هيأته لم تناسب "الصب" من الوقت.
Le commandant Slimane se plaçait toujours plus haut que ces attaques gratuites et orientées au point qu'il se prenait lui-même pour un «décalé» tellement tout lui paraissait décalé autour de lui.	قائد سليمان وضعت دائما فوق هذه الهجمات لا مبرر له والتوجه إلى النقطة التي تولى لنفسه "متداخلة" بحيث بدا تحول كل شيء من حوله	قائد سليمان يتواجد دائما فوق هذه الهجمات لا مبرر له والتوجه إلى النقطة التي تولى لنفسه "متداخلة" بحيث بدا تحول كل شيء من حوله
Et j'en veux pour preuve, ses pairs au sein du Conseil de la Révolution qui	وأريد أن أثبت أقرانه في مجلس الثورة يعرف كل أن الطفل من القديمة كان شيئا "خاصة" Tihert	وأريد أن أثبت أقرانه في مجلس الثورة القديمة Tihert يعرف كل أن الطفل من كان شيئا " خاصا " بالمقارنة مع البر

savaient tous que l'enfant de l'antique Tihert avait quelque chose de «spécial» par rapport à la bien-pensance de l'époque.	بالمقارنة مع البر الذاتي من الوقت.	الذاتي من الوقت
Moi, personnellement,	لي، شخصيا	انا شخصيا
Kaïd Ahmed était de cette race qui abhorrait de participer à une bagarre sans y prendre part d'une manière ou d'une autre.	وكان قايد أحمد هذا السباق الذي يمقت المشاركة في معركة من دون أخذ جزء في شكل او في آخر	وكان قايد أحمد هذا السباق الذي يمقت المشاركة في معركة من دون الانحياز في شكل او في آخر
Vous savez, je ne fais pas partie de ceux qui cherchent des poux dans la tête des autres.	أنت تعرف، أنا لست واحدا من أولئك الذين يسعون القمل في أذهان الآخرين.	أنت تعرف، أنا لست واحدا من أولئك الذين يبحثون القمل في رؤوس الآخرين.
Kaïd Ahmed est mort de son exil injuste. Il est parti parce qu'il a trop aimé sa patrie (sans jeu de mots) jusqu'à la mort, même si personne aujourd'hui n'est capable de dire avec certitude de quoi est mort cet homme si particulier.	توفي أحمد قايد غير عادلة في المنفى. غادر لأنه يحب بلده جدا (لا يقصد التوريه) حتى الموت، حتى لو لم يكن أحد اليوم يستطيع أن يقول على وجه اليقين ما هو خاص جدا حول هذا الرجل مات	توفي أحمد قايد غير عادلة في المنفى. غادر لأنه يحب موطنه جدا (لا يقصد التوريه) حتى الموت، حتى لو لم يكن أحد اليوم يستطيع أن يقول على وجه اليقين ما هو خاص جدا حول هذا الرجل مات
Les révolutions dans les pays voisins et autres ne sont pas des parties de plaisir	الثورات في الدول المجاورة وأجزاء أخرى ليست متعة	الثورات في الدول المجاورة و اطراف أخرى ليست متعة
y compris pour ceux qui réfléchissent à trois fois avant d'ouvrir leur bouche.	بما في ذلك أولئك الذين يعتقدون ثلاث مرات قبل فتح أفواههم	بما في ذلك أولئك الذين يفكرون ثلاث مرات قبل فتح أفواههم

on finit par laisser apparaître la grosse montagne de poussière cachée sous le tapis	ونحن في النهاية تكشف عن الجبل الكبير من التراب مخبأة تحت السجادة	ونحن في النهاية تكشف عن الجبل الكبير من الغبار مخبأة تحت السجادة

II-2- Article II

tout en prenant position pour les victimes de la sauvagerie coloniale dont les chantres cherchent aujourd'hui à se refaire une virginité et à nier ce moment sale de l'Histoire «occidentale	بينما اتخاذ موقف لضحايا الوحشية الاستعمارية التي المطربين يسعون الآن لإعادة بناء عذريتهن، وكذلك حرمان هذه اللحظة القذرة التاريخ الغربية"	بينما اتخاذ موقف لضحايا الوحشية الاستعمارية التي المطربين يسعون الآن لإعادة بناء عذريتهن، وكذلك انكار هذه اللحظة القذرة التاريخ الغربية"
On convoque l'Histoire pour répondre au discours colonial décrivant les Algériens	واستدعى التاريخ لتلبية الخطاب الاستعماري يصف الجزائريين	واستدعى التاريخ للرد علي الخطاب الاستعماري يصف الجزائريين
Ainsi, les premières pièces historiques s'inscrivaient dans ce nécessaire dialogue avec l'Autre,	وهكذا، كانت القطع التاريخية الأولى جزءا من هذا الحوار الضروري مع الآخر	وهكذا، كانت المسرحيات التاريخية الأولى جزءا من هذا الحوار الضروري مع الآخر
évitant les dénonciations directes et franches.	وتجنب الانسحاب المباشر والعلني.	وتجنب الكشف المباشر والعلني
Mais tout cela allait provoquer l'ire des autorités coloniales qui réagirent négativement en censurant les pièces et les auteurs.	. ولكن كل هذا يثير حفيظة السلطات الاستعمارية الذين كان رد فعل سلبيا الرقابة أجزاء من الكتاب و	. ولكن كل هذا يثير حفيظة السلطات الاستعمارية الذين كان رد فعل سلبيا الرقابة المسرحيات من الكتاب و
quelques auteurs algériens vont convoquer l'Histoire	فإن بعض الكتاب إلى التاريخ الجزائري عقد عزل المحتل الذي	فإن بعض الكتاب إلى التاريخ الجزائري عقد عزل المحتل الذي

pour mettre en accusation l'occupant qui détient tous les espaces de pouvoir.	يحمل جميع المساحات السلطة	يتولى جميع المساحات السلطة
Toute la période qui a précédé la guerre de libération a consisté, chez certains auteurs «indigènes»	طوال الفترة التي سبقت حرب التحرير كانت، في 'الأم بعض الكتاب	طوال الفترة التي سبقت حرب التحرير كانت، في المحليين بعض الكتاب
Le sens prend un glissement extraordinaire en épousant les contours d'un autre schéma politique et idéologique.	الاتجاه يأخذ زلة الخاصة ملامح مخطط آخر السياسية والأيديولوجية.	المعنى يأخذ زلة الخاصة ملامح مخطط آخر السياسية والأيديولوجية.
La suite théâtrale de Kateb Yacine,	كاتب ياسين بعد المسرح،	كاتب ياسين بعد المسرح،

Nous sommes tous d’accord que tous les signes linguistiques sont polysémiques, ils changent de sens selon leur position dans le cotexte où ils apparaissent, et aussi selon le contexte où ils s’inscrivent. Dans le tableau ci-dessus, où nous avons recueilli les unités qui ont été traduites par leurs équivalents dénotés ; par conséquent Google translate a aboutit à un contre-sens avéré étant donné que l’unité de l’énoncé source n’a pas obéi aux paramètres contextuels. Ceci dit, nous n’avons pas mentionné toutes les unités dont le sens peut changer d’un co(n)texte à un autre ; mais nous avons mis en lumière les unités polysémique dont le problème de déterminer le bon sens était un obstacle pour Google Traduction.

Ainsi, nous avons proposé dans la troisième colonne le sens approprié de toute unité polysémique mal traduite par Google. Nous remarquons que le nombre des mots polysémiques est vraiment important et nous essayerons d’expliquer, dans ce qui suit, l’effet de ces unités sur la construction du sens de l’énoncé.

III-Le sens d’un énoncé

Le sens de tout énoncé se construit par le résultat des sens de toutes les expressions qui le composent. Or, le sens de chaque expression, à son tour, dépend du sens global de l’énoncé. Ainsi, toute unité polysémique participe à l’élaboration des deux sens, c’est-à-dire le sens des expressions ainsi que le sens de l’énoncé.

Dans son ouvrage *Cours de linguistique générale* Ferdinand De Saussure explique davantage qu’ « *il ne suffit pas de considérer le rapport qui unit les diverses parties d’un*

syntagme entre elles… il faut tenir compte aussi de celui qui relie le tout et ses parties. »[76], c'est-à-dire qu'il y a un certain mécanisme dans la construction du sens global de l'énoncé. Déterminer le sens d'une unité dépend des autres unités, et chaque expression dépend, à son tour, des unités et aussi des autres expressions dans le même énoncé.

L'énoncé est une actualisation de la langue dont le sens dépend des relations qu'entretiennent les sens des expressions qui le composent. En outre, Emile Benveniste affirme dans le même point de vue qu' « *une phrase constitue un tout, qui ne se réduit pas à la somme de ses parties ; le sens inhérent à ce tout est réparti sur l'ensemble des constituants. Le mot est un constituant de la phrase, il en effectue la signification ; mais il n'apparait pas nécessairement dans la phrase avec le sens qu'il a comme unité autonome.* »[77].

Nous remarquons par ce passage qu'il y a un dynamisme dans la construction du sens de l'énoncé. Par ailleurs, les unités polysémiques jouent un rôle important dans la détermination du sens global de l'énoncé d'une part, et d'autre part cet énoncé joue un rôle important, à son tour, dans l'identification du sens de l'unité polysémique.

Si nous observons les unités polysémiques dans le tableau cité ci-dessus, nous comprendrons mieux ce dynamisme dans la construction du sens. Nous remarquons que les mots polysémiques traduites littéralement sur Google Traduction ont comme corollaire une destruction du sens des expressions données en arabe par Google.

Nous prenons, par exemple, l'expression « *Mais tout cela allait provoquer l'aire des autorités coloniales qui réagirent négativement en censurant les pièces et les auteurs.* ». Dans cette expression « pièce » est une unité polysémique qui peut regrouper une dizaine de sens différent. Google traduction a traduit cette unité polysémique comme suit : « ولكن كل هذا يثير حفيظة السلطات الاستعمارية الذين كان رد فعل سلبيا الرقابة أجزاء من الكتاب و. ». Google dans sa traduction a traduit le mot « pièce » par son sens propre c'est-à-dire son sens dénoté. Or, en se référent au sens global de l'expression et de l'énoncé en français nous comprenons qu'il s'agit d'une *pièce théâtrale.* De ce fait, l'expression donnée en arabe est dépourvue de tout sens.

Ainsi, tous les mots polysémiques mal traduits que nous avons cités ont donné des expressions qui n'ont aucun sens. En outre, l'ensemble de toutes ces unités ont donné, à leur tour, un énoncé dépourvu, également, de sens.

Nous concluons que le sens d'un énoncé est le résultat d'un double mouvement, puisque ce sens est en fonction du sens des expressions qui le composent, mais qu'inversement le sens de ces expressions dans cet énoncé est relatif au sens global de

---

[76] - Ferdinand De Saussure, Op.cit, p.172.
[77] - Emile Benveniste, *Problèmes de linguistique générale I*, PP.123-124.

l'énoncé lui-même. Guillaume affirme qu' « *une partie dans un tout est autre chose que cette partie isolée ou dans un autre tout* »[78]. Les interactions entre partie et tout, dans un énoncé, donnent à cet énoncé une forme cohérente. De ce fait, toute unité polysémique joue un rôle important pour que l'énoncé garde cette bonne forme comme nous le constatons lors de la lecture de l'article en français.

Dans l'énoncé, les unités linguistiques s'agencent de telle façon que chacune d'elles contribue à préciser le sens des autres de manière harmonieuse, donnant ainsi une cohérence au système. C'est cette cohérence qu'un traducteur doit garder lors de sa tentative de faire passer un message d'une langue à une autre en étant fidèle au sens globale de ce même énoncé.

Ainsi, pour J. Kristeva « *le texte est une productivité, ce qui veut dire que son rapport à la langue dans laquelle il se situe est redistributif.* »[79], c'est-à-dire que le texte est un espace non homogène et qu'il y a dans le texte lui-même une certaine production de sens, c'est-à-dire que le sens change au fur et à mesure que les unités qui le composent changent de position et de sens. Nous parlons ici d'un sens dynamique qui se construit avec la position des unités et des expressions dans le texte.

IV-Étranger et étrangeté selon Jean Claude Cauquet

Que nous traduisons d'une langue à une autre, nous passons d'un système obéissant à un rapport de sens entre les unités polysémiques à un système différent mais qui obéit, à son tour, aux mêmes principes. L'opération de traduction devient de plus en plus difficile quand nous avons affaire à deux langues qui ne partagent aucun lien historique, mais aussi aucun lien de la structure interne telles l'arabe et le français.

Ainsi, pour notre corpus, il s'agit de deux univers différents ainsi que de deux systèmes différents dans leur fonctionnement. Par ailleurs, identifier le sens de l'unité polysémique n'est pas une tâche aisée et cela peut nuire au sens global des articles quand il s'agit d'une traduction automatique.

Dans un colloque intitulé *traduction et discours littéraire*, Jean Claude Coquet parle de ce qu'il appelle par *étranger* et *étrangeté* dans le langage. Pour Coquet, il y a deux manières de traduire. En premier lieu, nous pouvons transposer le sens des unités d'une langue à une autre, c'est-à-dire nous ne plaçons pas l'unité dans tout l'énoncé mais nous la

---

[78] - P. Guillaume, *La psychologie de la forme*, Flammarion, 1979, p.23.
[79] - J. Kristeva, *Comment parler à la littérature*, *Tel Quel,* Seuil, paris, 1968, p.47.

prenons séparément. Cette traduction mène à ce que Coquet désigne par *étrangeté* du langage parce qu'elle perd le sens général de l'énoncé.

D'autre part, nous essayerons de construire le sens global de l'énoncé de la langue source dans la langue cible. Autrement dit, il s'agit ici d'une interprétation du texte dans sa totalité dans une autre langue en tenant compte de plusieurs facteurs linguistiques et extralinguistiques. Il s'agit dans cette traduction de chercher les équivalences entre les deux énoncés des deux langues. Ainsi, Cette traduction mène à ce que Coquet qualifie de *l'étranger* dans le langage.

Si nous lisons la traduction de l'article donnée par «*Google translate*» sans lire auparavant l'article en langue française, nous remarquerons que le sens global de l'article est ambigu. Nous verrons, également que certaines unités polysémiques ont été mal traduites. Ainsi, dans le deuxième chapitre de cette partie pratique nous allons présenter le processus de traduction automatique sur Google ; puis nous exposerons les différents problèmes rencontrés par ce moteur de recherche. Et en dernier, nous verrons si cette traduction des unités polysémiques des deux articles va mener à l'étranger et / ou à l'étrangeté du langage.

Chapitre II

D'une langue étrangère à l'étrangeté du langage

Dans ce deuxième chapitre de la partie pratique, nous essayerons d'expliquer, à l'aide de notre corpus, comment se déroule le processus de traduction automatique. Ainsi, les systèmes sont multiples et chaque logiciel de traduction a son propre fonctionnement. Pour notre étude, nous nous intéresserons au traducteur automatique de Google.

Au XXI ème siècle, « *Google translate* » se classe au sommet des traducteurs automatiques en ligne. L'efficacité de ses traductions se voit dès que nous cherchons à traduire un texte automatiquement. Or, comment *Google translate* arrive-t-il à traduire d'une manière satisfaisante des milliers d'énoncés ? Google arrive-t-il à traduire correctement toutes les unités ? Mais ce qui nous intéresse le plus, c'est Google face aux unités polysémiques.

I-La désambiguïsation sémantique

Anne Abeillé appelle par *désambiguïsation sémantique des mots*[80] l'opération qui consiste à associer une occurrence donnée d'un mot ambigu avec l'un des sens de ce mot. La désambiguïsation sémantique est nécessaire dans toutes les applications en traitement automatique du langage naturel, en l'occurrence en traduction automatique.

Cette opération de désambiguïsation sémantique des unités polysémiques s'effectue en deux étapes :

- ✓ Regrouper tous les sens de l'unité polysémique ;
- ✓ Procéder à l'étiquetage sémantique de chacune de ces occurrences.

La première étape a pour but de regrouper les différentes occurrences d'un mot polysémique en classes représentant chacune un sens. Quant à la seconde étape, elle a pour but d'assigner un sens à chaque classe crée précédemment.

II-Google traduction : comment ça marche ?

Contrairement aux multiples moteurs de traduction en ligne qui essaye d'inclure dans un ordinateur tout le lexique d'une langue ainsi que ses règles grammaticales, Google traduction adopte une approche différente. Ainsi, au lieu d'inculquer aux ordinateurs l'ensemble des règles d'une langue, Google met dans la base de données de ces ordinateurs des milliers de documents déjà traduits par des professionnels. Il s'agit de traduction parvenant de livres, d'organisations comme les Nations-Unis et le site Web du monde entier.

[80] - Anne Abeillé, *Sciences cognitives : Sémantique pour l'analyse de la linguistique à l'informatique*, Ed Masson, Paris, 1997, p.55.

Les ordinateurs analysent, par la suite, ces textes à la recherche du modèle statistiquement significatif, c'est-à-dire des schémas entre le texte original et la traduction qui ne peuvent être le fruit du hasard. Tout modèle ainsi détecté est ensuite réutilisé pour traduire des textes similaires, répétés des milliards de fois. Ainsi, ce processus permet aux ordinateurs d'emmagasiner des milliards de modèles et de gagner en efficacité. En revanche, la réalisation d'un tel moteur de traduction demande un grand projet de financement.

Par ailleurs, pour certaines langues, peu de documents traduits sont disponibles, c'est la raison pour laquelle la qualité des traductions données par *Google translate* peut varier d'une langue à une autre.

Des chercheurs d'IBM ont réalisé dans les années 1980 des logiciels qui exploitent des principes mathématiques. Ainsi, Google a crée, à l'aide de ces logiciels, un outil de traduction automatique différent de tous les autres systèmes.

A l'opposé des autres systèmes de traduction qui s'intéressent à amener l'ordinateur à ce qu'il soit capable de reconnaitre le bon sens, *Google Translate* ne cherche pas à traiter une expression linguistique ; mais il la considère comme un texte qui a sans doute été déjà traduit dans le passé. Ainsi, le logiciel utilise une puissance de calcul pour rechercher cette même expression dans un texte déjà existant et couplé à sa traduction.

Pour se faire, Google enregistre dans des bases de données des corpus immenses regroupés par ces différentes sources :

- ✓ Les documents émis par l'UE depuis 1957 en deux douzaines de langues;
- ✓ Tout ce que les Nations Unies et ses institutions ont pu mettre par écrit en six langues officielles ;
- ✓ D'énormes masses d'autres matériaux depuis les archives des tribunaux internationaux et les rapports de grandes sociétés ;
- ✓ Tous les articles et les ouvrages bilingues qui on été mis en ligne par des individus, des bibliothèques, des auteurs, et des départements universitaires.

Ainsi, d'énormes quantités de textes sont mises en parallèle avec leurs traductions dans des bases de données. Nous désignons ces corpus en traitement automatique du langage naturel par : « *Corpus bilingue ou multilingue alignés* »[81].

En outre, pour trouver la traduction la plus adéquate de l'expression à traduire, Google Traduction applique des méthodes statistiques[82] sur l'ensemble des correspondances établies

[81] - F. Rastier, *Représentation du contenu lexical et formalismes de l'Intelligence Artificielle*, Langages, 1987, p.22.

sur ces millions de paires de documents. Ainsi, David Bellos affirme dans son ouvrage *Le poisson et le bananier* : *Une histoire fabuleuse de la traduction* que cette méthode utilisée par Google traduction *« marche assez souvent… C'est tout à fait satisfaisant »* et que ce résultat *« explique pour une bonne part, le retour en grâce de l'utopie fondatrice de Warren Weaver : la perspective de réaliser un jour la FAHQT »* c'est-à-dire la réalisation de la traduction de haute qualité complètement automatisée.

A l'heure actuelle, *Google translate* offre des possibilités de traduction réversible entre 58 langues, soit 3 306 services de traductions distincts. Par ailleurs, Google traduit, également, des textes de deux langues qui ne partagent aucun lien historique en matière de traduction. Nous citons par exemple les relations traductionnelles :

Islandais → farsi, yiddish → vietnamien.

La qualité de traduction donnée par *Google translate* peut varier d'une langue à une autre. Cette différence dans la qualité de la traduction s'explique par les fortes variations de taille des corpus linguistiques accessibles à l'exploration de « *Google Translate* ». Par ailleurs, le logiciel tire aussi profit des capacités des traducteurs humains, car il ne manque jamais de demander à ses utilisateurs de suggérer une meilleure traduction que celle qu'il fournit.

II-1-Calcul de sens polysémique sur Google traduction

Nous avons vu ci-dessus que *Google translate* adopte la méthode statistique dans ses traductions. Pour se faire, nous avons vu que d'immenses quantités de corpus sont enregistrés dans des bases de données. Nous allons essayer dans ce qui suit de présenter le processus de traduction sur Google selon la méthode statistique de corpus parallèle. Or, nous signalons que nous n'avons pas les chiffres exacts utilisés par *Google Translate* parce que c'est un logiciel qui appartient à l'entreprise Google. Il est à la fois confidentiel et coûte des milliards de dollars, ce qui dépasse les moyens que nous avons pour la réalisation de ce travail.

II-2-La méthode statistique

Warren Weaver était le premier à avoir l'idée d'appliquer des méthodes statistiques à la traduction en 1949. Or, des limitations dues à la puissance des ordinateurs, disponibles en ce temps là, ont mis à l'échec cette approche dite statistique.

Par la suite, l'approche statistique refait surface au début des années 1990 grâce aux ordinateurs de plus en plus puissants qui se développent d'une part, et la disposition d'immenses corpus bilingue alignés d'autre part. En outre, la publication de groupe de

---

[82] - David Bellos, Op.cit, p.172.

chercheurs du Watson Research Center d'IBM sur *les fondements mathématiques de l'approche statistique*[83] a grandement contribué au regain d'efforts mis dans cette approche.

Pour traduire ses textes, Google traduction adopte l'approche statistique dont les composantes suivantes sont nécessaires :

1- Un corpus bilingue dont les phrases sont alignées ;
2- Un modèle probabiliste du langage ;
3- Le concept d'alignement de mots ;
4- Un modèle probabiliste de traduction ;
5- Un algorithme de recherche de traductions appelé décodeur.

Nous essayerons d'expliquer dans ce qui suit les deux premières composantes. Quant aux autres composantes, nous ne pouvons pas les entamer dans notre travail, parce qu'en premier lieu, nous sommes dans le domaine de la linguistique, donc, nous n'allons pas trop nous étaler dans l'application de traduction automatique. Dans un second lieu, après l'étude des cinq composantes, nous avons jugé que les deux premières seront suffisantes pour comprendre le processus de calcul de sens des unités polysémiques ; et ceci nous permettra d'atteindre notre objectif tracé dès le départ.

II-2-1-Corpus parallèles

Un corpus parallèle est un élément important en traduction automatique statistique sur Google. Ce corpus consiste en une immense base d'exemples de traductions qui permet d'avoir un modèle probabiliste du langage cible et le modèle probabiliste de traduction. Autrement dit, un corpus parallèle est un ensemble de phrases S c'est-à-dire de langue source ; et un ensemble de phrases C, c'est-à-dire de langue cible où l'ensemble des phrases de langue C sont considérées comme des traductions des phrases de langue S.

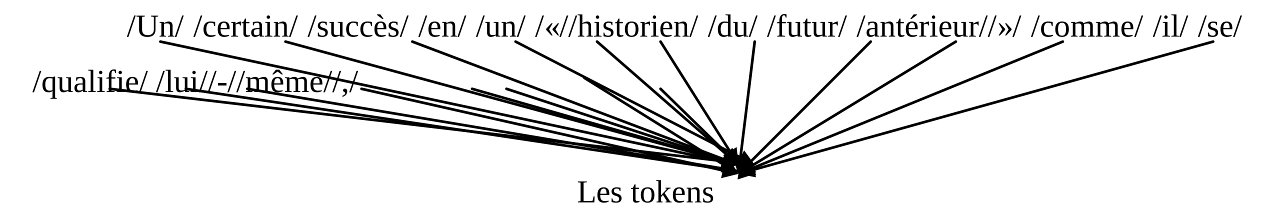
Pour mieux expliquer ce que nous venons d'avancer, nous prenons une phrase du premier article, par exemple le titre : «Si les hommes sont oublieux, l'histoire a bonne mémoire ». Cette phrase de langue Source (S), qui est le français, existe sans doute dans un corpus, dans la base de données de *Google translate* avec la phrase de la langue Cible (C) « **إذا كان الرجال هم النسيان، قصة لديه ذاكرة جيدة** » qui existe à son tour dans un autre corpus mis en parallèle au corpus de la langue source. Nous signalons que cette phrase peut ne pas exister sous cette structures exactement, mais chaque unité qui la compose a sa traduction en parallèle dans la base de données de Google traduction.

---

[83] - Y.M. Visetti., *Des systèmes experts aux systèmes à bases de connaissances : à la recherche d'un nouveau schéma régulier*, Intellectica, 1991, p. 16.

A l'aide d'un ensemble de phrases existant dans des corpus parallèles dans la base de données de *Google translate*, Google arrive au moment de la traduction à déduire une distribution de probabilité pour le modèle de traduction.

Pour tout calcul de sens dans la méthode statistique sur *Google translate*, le logiciel commence son opération par la *tokénisation*[84]. Par *tokénisation*, nous entendons l'opération qui consiste à segmenter une phrase en objet appelés tokens. Ce dernier désigne soit un mot, soit un signe de ponctuation, un guillemet, une parenthèse ou un nombre. Il est l'élément fondamental sur lequel repose la traduction automatique statistique. Pour plus de clarté, puisque nous sommes dans le domaine de la linguistique, la tokénisation est ce que nous appelons en linguistique la segmentation et les tokens ce sont les unités minimales significatives et non significatives.



L'opération de tokénisation sur Google traduction

II-2-2-L'opération de traduction

Afin de traduire une phrase s de la langue source S, nous supposons que toutes les phrases c de la langue cible C sont des traductions possibles de s. nous schématisons donc la traduction statistique par le biais du modèle d'un canal de communication bruité. Nous notons donc Pr (c/s), la probabilité conditionnelle que la phrase cible c soit une traduction de la phrase source s. Pour calculer la phrase cible, la méthode statistique utilise *la formule de Bayes*[85] qui se résume comme suit :

$$Pr(c/s) = \frac{Pr(c)\ Pr(s/c)}{Pr(s)}$$

- ✓ Pr (c/s) désigne la probabilité conditionnelle que la phrase c de langue cible C qui est l'arabe soit une traduction de la phrase s.

[84] - Anne Abeillé, *Op.cit.* p.70
[85]- Visetti Y.M. Op.cit, p.22.

Ainsi, puisque s est fixe et indépendante de c, ce problème de maximisation, que Visetti appelle *le problème fondamental de la traduction automatique statistique*[86], devient :

$$\hat{c} = \text{argmax } Pr(c)Pr(s/c)$$

- ✓ Pr(c) désigne le modèle probabiliste du langage cible.

Où

- ✓ Pr (s/c) désigne le modèle probabiliste de traduction.
- ✓ Argmaxe désigne le problème de recherche.

Nous cherchons ĉ l'une des phrases cibles qui maximise Pr (c/s) par le biais de la formule de Bayes.

Nous prenons pour exemple une phrase tirée de notre corpus : « *c'est du moins là l'espoir que nourrissent les jeunes* ». Cette phrase s de langue source S qui est le français sera calculée avec la formule cité ci-dessus.

Pour notre étude sur la polysémie, nous expliquerons le processus de calcul du sens polysémique « jeunes » dans cette phrase. Ainsi, les traductions possibles de l'unité « jeunes » dans la base de données de *Google translate* sont :

- ➢ الصغار
- ➢ الشباب

Donc, nous avons deux probabilités conditionnelles :

- ➢ Pr (الصغار/jeunes)
- ➢ Pr (الشباب/jeunes)

Après l'application de la formule de Bayes sur les deux probabilités Google choisit celle qui maximise. Donc le calcule sera comme suit :

Pr (الصغار) * Pr (jeunes/الصغار)

Pr (الشباب) * Pr (jeunes/ الشباب)

Le calcul des phrases et des mots selon cette formule permettra de dégager le sens des mots polysémiques par rapport au contexte où il apparait dans tous les corpus enregistrés dans la base de données de *Google translate*.

---

[86] - Ibid.

II-2-3-Modèle probabiliste du langage

Pour comprendre le sens d'une unité polysémique dans une phrase, l'être humain fait recours au sens des unités qui entourent cette unité, nous tenons compte également du sens global de l'énoncé. Par ailleurs, il est difficile parfois de tenir compte du sens d'une unité lexicale par rapport au contexte de tout l'énoncé, donc nous nous limitons à la seule phrase où ce mot apparait afin de dégager le sens connoté. Ainsi, Saussure lui-même affirme qu'une unité n'a de valeur dans une phrase que par rapport aux autres unités de cette même phrase.

Pour la désambiguisation sémantique d'une unité polysémique, *Google translate* se sert du même principe. Autrement dit, pour dégager le sens d'un mot polysémique, il calcule le sens des autres unités qui le précède. Ainsi, pour calculer *Google translate* se base sur l'opération suivante : la probabilité conditionnelle d'occurrence d'un mot wn ne dépend que des n-1 mots qui le précèdent soit :

Pr (wn/w1,…, w-1)

Ainsi, pour calculer le sens d'une unité polysémique, Google traduction tient compte d'un nombre de mots qui précèdent cette unité. Mais nous ignorons le nombre exact utilisé parce que ce logiciel de traduction est confidentiel. Ainsi, pour comprendre l'opération de calcul du sens, nous calculons le sens d'une unité polysémique de notre corpus en utilisant le nombre trois par exemple.

Nous prenons toujours la phrase « *c'est du moins là l'espoir que nourrissent les jeunes* ». Nous ferons ici un calcul de sens par la segmentation des unités avec un nombre de 3.

Ainsi :

3 grammes		
1 gramme	**2 grammes**	**3 grammes**
C'	est	du
est	du	moins
du	moins	là
moins	là	l'
là	l'	espoir
l'	espoir	que
espoir	que	nourrissent
que	nourrissent	les

nourrissent	les	jeunes
les	jeunes	.

Après la segmentation de la phrase selon le nombre de 3 grammes, Google passe aux calcule des probabilités comme suit :

$$Pr\ (w^{12}) = \prod_{i=1}^{12} (Pr\ \ wi\ /\ wi\text{-}2,\ wi\text{-}1)$$

- ✓ 12 indique le nombre de mots que nous avons dans notre phrase.

Donc nous aurons :

Pr (w^{12}) = Pr (c / 0, 1) * Pr (est / c, 0) * Pr (du / est, c) * Pr (moins / du, est) *
Pr (là / moins, du) * pr (l' / là, moins) * Pr (espoir / l', là) *
Pr (que / espoir, là) * Pr (nourrissent / que, espoir) *
Pr (les / nourrissent, que) * Pr (jeunes / les, nourrissent).

Nous signalons que chaque unité a une valeur que nous ignorons, pour cela nous ne pouvons pas faire les calculs nous-mêmes, mais nous avons juste présenté la méthode suivie lors de l'opération traduisant sur *Google translate*.

Ainsi, nous avons par exemple l'unité polysémique « nourrissent » qui peut être traduite soit par :

- ✓ اطعام

Où

- ✓ تغذية

Après l'opération de segmentations et le calcul des probabilités, l'unité « nourrissent » est calculée selon son occurrence avec les unités « espoir », « que », « les », « jeunes ».ainsi, nous savons que même un être humain va se servir de ces unités pour la désambiguïsation sémantique de l'unité « nourrissent ».

III-Problèmes de traduction sur Google

Après la présentation du processus de désambiguïsation sémantique des unités polysémique en traduction automatique sur Google, nous allons passer à l'explication des problèmes rencontrés par Google lors de sa traduction de notre corpus.

Lors de l'opération de traduction, Google se trouve face à de différents obstacles qui rendent sa traduction des textes mauvaise. Nous allons donner quelques exemples que nous

avons extraits de notre corpus et expliquer les raisons pour lesquelles les unités polysémiques ont été mal traduites.

III-1-Problèmes liés au contexte dans le processus traductique

Nous avons présenté dans le chapitre précédent le rôle que joue le contexte dans la désambiguïsation sémantique des unités polysémique. Or, nous ne pouvons pas avoir toutes les connaissances du monde extralinguistique. Ainsi, les unités polysémiques acquièrent leur sens au moment de l'énonciation.

Dans l'exemple « c'est du moins là l'espoir que nourrissent les jeunes », « nourrissent » nécessite des connaissances sur le contexte où l'énoncé a été dit. Par ailleurs, Google a mal traduit « nourrissent » par son sens dénoté « إطعام », alors que « nourrissent dans ce passage a été employé au sens connoté. Le système de traduction de Google ne possède pas des connaissances qui lui permettront de savoir que « espoir » est une chose virtuelle qui ne peut pas être nourri dans le sens de « manger ».

Par ailleurs, Google ne garde pas aussi l'enchainement des sens de tout l'énoncé dans un contexte donné. Ainsi, dans l'expression « Mais tout cela allait provoquer l'ire des autorités coloniales qui réagirent négativement en censurant les pièces et les auteurs » nous n'avons pas eu la bonne traduction de « pièces » qui a été aussi traduit au sens dénoté parce que Google n'a pas gardé le sens générale où s'inscrit le contexte de l'énoncé pour savoir qu'il s'agissait d'une pièce théâtrale.

III-2-Problèmes liés au cotexte dans le processus traductique

La position qu'occupe l'unité polysémique dans le système de la langue joue un rôle important dans la désambiguïsation sémantique de cette même unité. Ainsi, lors de la segmentation des unités, Google n'arrive pas, dans beaucoup de cas, à déterminer la position exacte de l'unité polysémique dans le système. Ainsi, l'expression « *Et j'en veux pour preuve, ses pairs au sein du Conseil de la Révolution qui savaient tous que l'enfant de l'antique Tihert avait quelque chose de «spécial» par rapport à la bien-pensance de l'époque.* » a été traduite par :

وأريد أن أثبت أقرانه في مجلس الثورة يعرف كل القديمة كان شيئا "خاصة" Tihert بالمقارنة مع البر الذاتي من الوقت. أن الطفل من

Nous avons dans ce passage le mot polysémique « spécial » qui peut être traduit soit par « خاصة » ou « خاصا ». Ainsi, dans la phrase de notre corpus « spécial » a été traduit par « خاصة » et il a donné à l'expression un sens autre que celui exprimé dans l'énoncé original.

Nous expliquons l'échec de la traduction de ce mot polysémique dans cette expression par trois réponses. En premier lieu, parce que Google lors de la segmentation des unités afin de calculer le sens ; l'unité « spécial » a été calculée par rapport aux guillemets au lieu de calculer son sens par rapport à « Tihert » qui se positionne un peu loin de « spécial ».

En second lieu, l'unité polysémique a été mal traduite à cause de la longueur de l'énoncé, ce qui ne facilite pas la tâche de la segmentation des unités pour Google. Ainsi, si nous allons traduire juste la phrase « *Tihert avait quelque chose de spécial* », « spécial » sera traduit par « خاصا ».

Et dans un dernier lieu, « Tihert » est un nom propre qui ne se trouve pas dans la base de données de Google, c'est-à-dire, il n'a pas sa traduction en arabe. Ainsi, nous savons que le système d'écriture du français commence de gauche vers la droite, en revanche le système arabe commence de la droite vers la gauche. De ce fait l'unité « Tihert » qui ne peut pas être traduite par Google garde toujours la position gauche dans le système, quant aux autres unités traduites, elles prennent la position droite.

En outre, nous avons un autre problème qui ne facilite pas la traduction des unités polysémiques pour *Google translate* ; ce problème se situe au niveau de la ponctuation. Ainsi, dans le passage « Moi, personnellement, » traduite par « لي، شخصيا » l'unité polysémique « moi » a été mal traduite pour une seule raison : la présence de la virgule après l'unité « moi » qui a séparé cette unité de l'unité « personnellement ». De ce fait Google l'a traduite séparément ; ce qui a donné une traduction erronée. Or si nous essayerons de traduire ce même passage mais en relevant la virgule, nous aurons une traduction correcte « انا شخصيا ».

Enfin, les unités polysémiques ne doivent pas être constituées d'une liste exhaustive de sens, mais nous devons prendre en considération leur mode d'emploi dans le système linguistique. Nous appelons cette, en traitement automatique du langage naturel, *la compositionalité*[87]. Cette dernière consiste à prendre en considération la relation entre l'analyse syntaxique et l'analyse sémantique. Autrement dit, c'est le fait de prendre en considération les relations entre la syntaxe et la sémantique parce qu'il existe un couplage entre ces deux parties de la langue.

III-3-D'une langue étrangère à l'étrangeté du langage

Nous sommes arrivés à la dernière partie de notre travail dans laquelle nous allons voir si nous avons atteint le but que nous nous sommes fixé au départ.

---

[87] - Y.M. Visetti. Op.cit. p. 34.

Nous avons vu comment Google a rencontré des difficultés à traduire les unités polysémiques ; et ceci a donné des expressions en arabe dépourvues de tout sens. De ce fait, les unités polysémiques des deux articles, de notre corpus, que nous avons traduites sur Google ont été mal traduites.

Ainsi, la lecture des deux articles en arabe est une preuve de traduction de l'étrangeté selon les termes de Jean Claude Coquet. Il s'agit d'une traduction mot-à-mot et non du sens global de l'énoncé. Par cela nous concluons que les unités polysémiques jouent un rôle important dans la construction du sens dans toutes les langues, et jusqu'à nos jours les chercheurs en TALN n'ont pas pu résoudre les problèmes de polysémie.

Synthèse

Nous avons vu dans cette partie pratique comment les unités polysémiques changent de sens et passe de la signification au sens. Elle constitue un élément important dans toutes les langues et joue un rôle primordial dans la construction du sens des énoncés.

Nous avons présenté, par la suite, le processus de traduction automatique sur Google translate, ainsi que le la méthode de désambiguïsation sémantique des unités polysémiques selon la méthode statistique.

Des corpus de traductions, regroupés par différentes sources, sont mis en parallèle dans des bases de données ; et Google se sert de sa capacité de calcul afin de trouver la traduction la plus adéquate.

Or, la difficulté de segmenter les unités dans les deux articles n'a pas permis de donner à chaque unité polysémique sa position exacte lors de la traduction. En outre, des connaissances extralinguistiques manquaient à Google pour qu'il puisse déterminer le sens dénoté des sens connotés des unités. Ainsi, toutes ces difficultés ont mené à ce que Jean Claude Coquet désignait par la traduction de l'étrangeté dans le langage.

Malgré ces échec, Google translate arrive à donner une aide à l'être humain. Ainsi, il permet de traduire des documents scientifiques qui et ceci facilitera le travail de grandes entreprises.

Nous concluons par la présentation des deux procédures de calcul de sens sur Google translate, à savoir que ce dernier utilise les mêmes principes qu'un être-humain. Si l'Homme se sert de son cerveau pour une preuve d'intelligence ; l'ordinateur se sert de ses calculs pour que nous puissions parler aujourd'hui d'intelligence artificielle.

Synthèse

Nous avons vu dans cette partie pratique comment les unités polysémiques changent de sens et passe de la signification au sens. Elle constitue un élément important dans toutes les langues et joue un rôle primordial dans la construction du sens des énoncés.

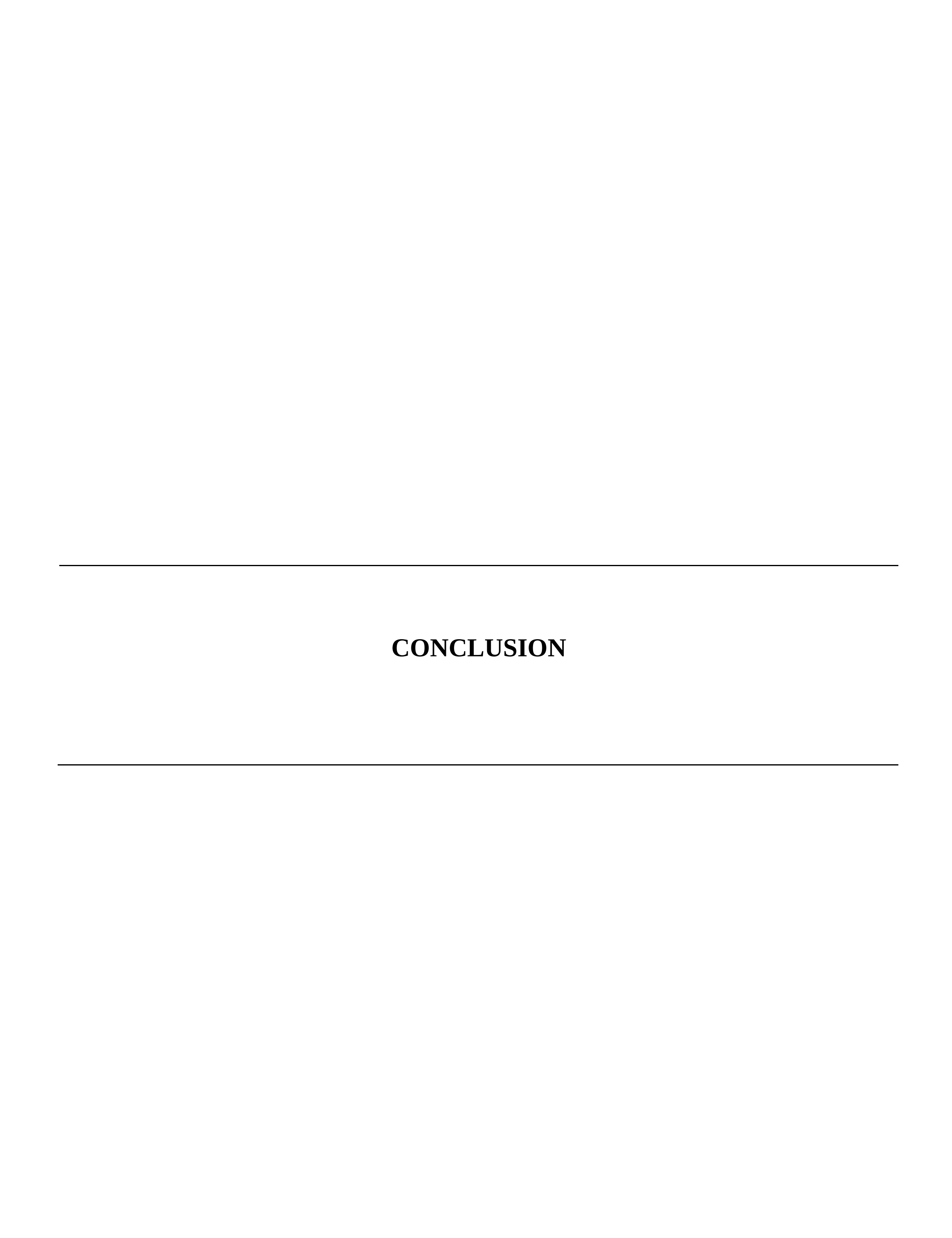
Nous avons présenté, par la suite, le processus de traduction automatique sur Google translate, ainsi que le la méthode de désambiguïsation sémantique des unités polysémiques selon la méthode statistique.

Des corpus de traductions, regroupés par différentes sources, sont mis en parallèle dans des bases de données ; et Google se sert de sa capacité de calcul afin de trouver la traduction la plus adéquate.

Or, la difficulté de segmenter les unités dans les deux articles n'a pas permis de donner à chaque unité polysémique sa position exacte lors de la traduction. En outre, des connaissances extralinguistiques manquaient à Google pour qu'il puisse déterminer le sens dénoté des sens connotés des unités. Ainsi, toutes ces difficultés ont mené à ce que Jean Claude Coquet désignait par la traduction de l'étrangeté dans le langage.

Malgré ces échec, Google translate arrive à donner une aide à l'être humain. Ainsi, il permet de traduire des documents scientifiques qui et ceci facilitera le travail de grandes entreprises.

Nous concluons par la présentation des deux procédures de calcul de sens sur Google translate, à savoir que ce dernier utilise les mêmes principes qu'un être-humain. Si l'Homme se sert de son cerveau pour une preuve d'intelligence ; l'ordinateur se sert de ses calculs pour que nous puissions parler aujourd'hui d'intelligence artificielle.



CONCLUSION

Au terme de notre recherche, il est d'ores et déjà utile de signaler que le langage auquel aboutit Google translate est à la fois une « étrangeté » mais parfois même étrange. Et ceci pour plusieurs raisons. Avant toutefois d'en citer quelques unes, rappelons d'abord l'utilité du langage dans la communication ordinaire (sociale).

Du besoin de communiquer, de la nécessité vitale de secours mutuels est né cet instrument de communication : la langue. Eminemment sociale et socialisante, elle permet de tisser et d'établir des liens sociaux.

Ainsi, dès le début des études sur les langues, les linguistes se préoccupaient des problèmes liés au sens et des phénomènes d'ambiguïté des énoncés. Ils le sont davantage aujourd'hui avec la naissance d'une discipline nouvelle : le Traitement Automatique du Langage Naturel (TALN).

Le traitement automatique du langage naturel est un domaine pluridisciplinaire qui regroupe la linguistique, l'informatique, l'intelligence artificielle et d'autres disciplines. Ce domaine nouveau en linguistique se fixe pour objectif l'automatisation du langage naturel, c'est-à-dire modéliser et reproduire, à l'aide des ordinateurs, le langage naturel. En effet, les chercheurs en Intelligence Artificielle (IA) étaient les premiers à avoir l'idée d'automatiser cette faculté propre à l'homme; et les premières définitions de l'IA firent de l'intelligence humaine l'objectif à atteindre.

Au XXI siècle, les techniques d'information et de communication (TIC) ne cessent de se développer de plus en plus, et des logiciels traitant les langues naturelles ont été crées. En revanche, ces derniers ne rendent pas toujours compte de manière satisfaisante du sens des énoncés à cause des ambiguïtés des unités inhérentes aux langues.

Dans notre étude, nous avons choisi de travailler sur la problématique de l'ambigüité des unités polysémique dans le domaine du traitement automatique du langage naturel. Ainsi, ce nouveau domaine regroupe plusieurs applications telles que la conversation homme-machine, le résumé automatique et la traduction automatique.

En effet, nous savons que l'ambiguïté des unités appartenant aux langues différentes se voit dès que nous passons d'une langue à une autre. Chaque langue divise son espace sémantique selon sa vision du monde, parce que les unités linguistiques sont en étroit rapport avec le domaine extralinguistique. Par ailleurs, chaque langue a, aussi, son propre système

auquel obéissent toutes ses unités. Pour toutes ces raisons, nous avons choisi d'étudier la problématique de la polysémie en TALN par le biais de la traduction automatique.

Ainsi, la traduction automatique était le premier domaine du TALN, elle remonte à la guerre froide entre les Russes et les Américains. Ces derniers se souciaient de leur ennemi, et ils ont eu l'idée de créer une machine qui traduira des phrases russes pour ses soldats. Or, les traductions ne comptaient qu'un nombre limité de phrases. Par la suite, manipuler des dictionnaires électroniques était les principaux travaux tracés en TALN.

Aujourd'hui, le traitement automatique du langage naturel devient un domaine très actif et de nombreuses applications commencent à atteindre le grand public, tels les résumés automatiques, la correction orthographique qui se fait par ordinateur ainsi que les divers logiciels de traduction en ligne parmi lesquels Google occupe la place la plus importante.

Google translate est un logiciel de traduction gratuit en ligne qui appartient à l'entreprise Google en Californie. Aujourd'hui, il nous permet de traduire instantanément des phrases, des documents, et même des sites web entiers en 58 langues.

En effet, si nous avons effectué notre travail sur Google translate et non pas sur un autre moteur de traduction, c'est parce que Google donne des résultats plus satisfaisants par rapport aux autres systèmes ; c'est la raison pour laquelle il occupe aujourd'hui le sommet de tous les traducteurs automatiques.

La polysémie est un fait linguistique, une partie importante dans toutes les langues naturelles. C'est une unité linguistique regroupant des sens connotés, mais liés à un seul sens dénoté. Ainsi, de nombreux travaux ont été effectués en linguistique pour traiter le problème du sens dans les unités polysémiques. Aujourd'hui, la polysémie constitue un sujet épineux en traitement automatique du langage naturel, en l'occurrence en traduction automatique.

Dans notre travail, nous avons essayé de présenter le problème de la polysémie en traduction automatique sur Google translate afin de répondre aux questions de notre problématique posées au départ.

Par ailleurs, les unités polysémiques changent de sens selon le contexte et le cotexte ; et passent de la signification au sens. Ainsi, la signification est le sens dénoté de l'unité polysémique ; quant au sens, c'est ce qui s'ajoute à ce premier sens dénoté. De ce fait, les

unités polysémiques changent d'un énoncé à un autre. Ainsi, ce changement mène à un dynamisme dans les énoncés ce qui fait que les langues évoluent d'une époque à une autre.

Pour enlever toute ambigüité liée au sens dans les unités polysémiques, nous avons recours au contexte où s'inscrivent ces unités et au cotexte dont lequel elles prennent position. En effet, Google translate se base sur les mêmes principes pour la désambiguïsation sémantique des unités polysémiques lors du processus de traduction.

Ainsi, au lieu d'inculquer tout le lexique d'une langue et ses règles grammaticales, Google adopte une méthode qui se base sur des traductions déjà faites par des traducteurs humains. Il adopte la méthode statistique du corpus parallèle. Ainsi, des milliers de documents déjà traduits sont mis en parallèle dans des bases de données. Par la suite, les ordinateurs analysent ces textes à la recherche du modèle statistiquement significatif, c'est-à-dire la recherche des schémas entre le texte de la langue source et le texte de la langue cible.

De ce fait, à l'opposé des autres systèmes de traduction qui cherchent à mener l'ordinateur à ce qu'il soit capable de reconnaitre le bon sens, Google translate ne cherche pas à traiter un énoncé ou une expression linguistique ; mais il utilise une puissance de calcul pour trouver cette même expression ou ce même énoncé dans les corpus disponibles dans la base de données de ses ordinateurs.

Dans un premiers temps, une segmentation des unités est nécessaire avant d'entamer le calcul des sens. Puis, Google translate calcule le sens d'une unité polysémique selon son occurrence dans tous les corpus où elle apparait en utilisant la formule de *Bayes*. Cette opération permet de dégager la probabilité conditionnelle qui maximise la traduction de cette unité.

Par la suite, Google passe à une autre opération où il calcule le sens de l'unité polysémique en tenant compte de sa position dans l'énoncé où elle apparait. Autrement dit, il calcule le sens de cette unité polysémique avec les autres unités qui l'entourent.

En revanche, lors du calcul du sens polysémique, Google translate se trouve face à de différents obstacles liés à l'ambigüité du sens des unités polysémiques. Nous citrons ces difficultés en essayant de répondre aux questions que nous avons posées au départ :

✓ Google translate ne possède pas des connaissances qui lui permettront de traduire d'une manière correcte les unités polysémiques. Ainsi, pour reprendre l'exemple de

l'expression « l'espoir que nourrissent les jeunes », des connaissances sont nécessaires pour comprendre que l'unité polysémique « nourrissent » a été utilisée en sens connoté et non au sens dénoté ; parce que « l'espoir » est une chose qui ne peut pas manger de la nourriture.

- ✓ La langue est dynamique et ses unités qui la composent sont en perpétuel évolution. Les unités polysémiques acquièrent de nouveau sens d'un contexte à un autre. ainsi, « pièce » signifie dans un premier sens, une partie ou un élément d'un tout ; mais aussi plusieurs connotations se sont rajoutées à ce premier sens au fil du temps selon différentes situations. De ce fait, nous ne pouvons pas regrouper toutes les connotations d'une unité polysémique.

- ✓ le sens de l'énoncé se construit par le résultat des sens de toutes les expressions qui le composent. En outre, le sens des expressions se construit, à son tour par le résultat des sens des unités qui le composent. Nous avons un dynamisme dans la construction du sens et les unités polysémiques jouent un rôle important dans la construction des deux sens ; c'est-à-dire le sens des expressions et le sens de l'énoncé. En outre, le sens de l'unité polysémique se construit, à son tour, par le sens global de l'énoncé ; c'est ce que Google translate n'arrive pas à prendre en considération et tombe dans des erreurs de traductions.

- ✓ Les unités polysémiques acquièrent leur sens dans le système de la langue. Ainsi, chaque unité se définit par rapport aux autres unités qui l'entourent. Or, lors de l'opération de segmentation, Google translate n'arrive pas à déterminer la position de chaque unité ; c'est-à-dire, il ne trouve pas le début et la fin de chaque unité.

- ✓ Enfin, chaque langue a sa propre syntaxe ; autrement dit, chaque langue a son propre fonctionnement. Pour l'étude de notre corpus, nous avons choisi de traduire les deux articles du français vers l'arabe. Nous avons donc, deux langues complètement différentes dans leur syntaxe. Ainsi, le système d'écriture en français commence de la gauche vers la droite, or le système de la langue arabe commence de la droite vers la gauche. En outre, les unités dans le système de la langue française et celui de la langue arabe ne prennent pas la même position. ainsi, si le verbe en français prend la deuxième position après le sujet ; en

arabe il prend la première position dans l'agencement syntaxique. Pour toute ces raisons Google translate se trouve face à de différents obstacles qui lui posent problème parce qu'il doit refaire tout le fonctionnement de l'énoncé.

Tous ces points que nous venons de citer ont été les problèmes étudiés en traitement automatique du langage naturel. Or jusqu'à nos jours, malgré les grands progrès qu'ils ont pu faire ; ces problèmes du traitement automatique de la langue n'arrivent pas à être résolus. Ainsi, lors de la traduction des deux articles de notre corpus nous avons eu un texte en arabe dépourvu de tout sens. Nous avons eu une traduction mot-à-mot ; nous n'avons pas eu une traduction du sens global de l'énoncé ou, selon les termes de Jean Claude Coquet une traduction de l'étrangeté ; mais nous avons eu comme résultat une traduction de la langue, une traduction de l'étrangeté.

Nous pouvons conclure que les unités polysémiques sont un fait important dans toutes les langues. Chaque unité polysémique participe à la construction du sens de l'énoncé. Ainsi le bon sens des unités polysémiques, dans les deux articles en langue française, a donné aux deux textes une bonne forme et un sens. Or la mauvaise traduction des unités polysémiques a comme corolaire deux textes dépourvus de tout sens accessible, deux textes qui mènent le lecteur à tomber dans la traduction de l' « étrangeté ».

Au fil de l'histoire du traitement automatique du langage naturel des progrès considérables ont été marqués. Nous somme passés des dictionnaires bilingues à des logiciels de traduction plus avancés en ligne. En revanche le problème du sens reste toujours difficile à résoudre.

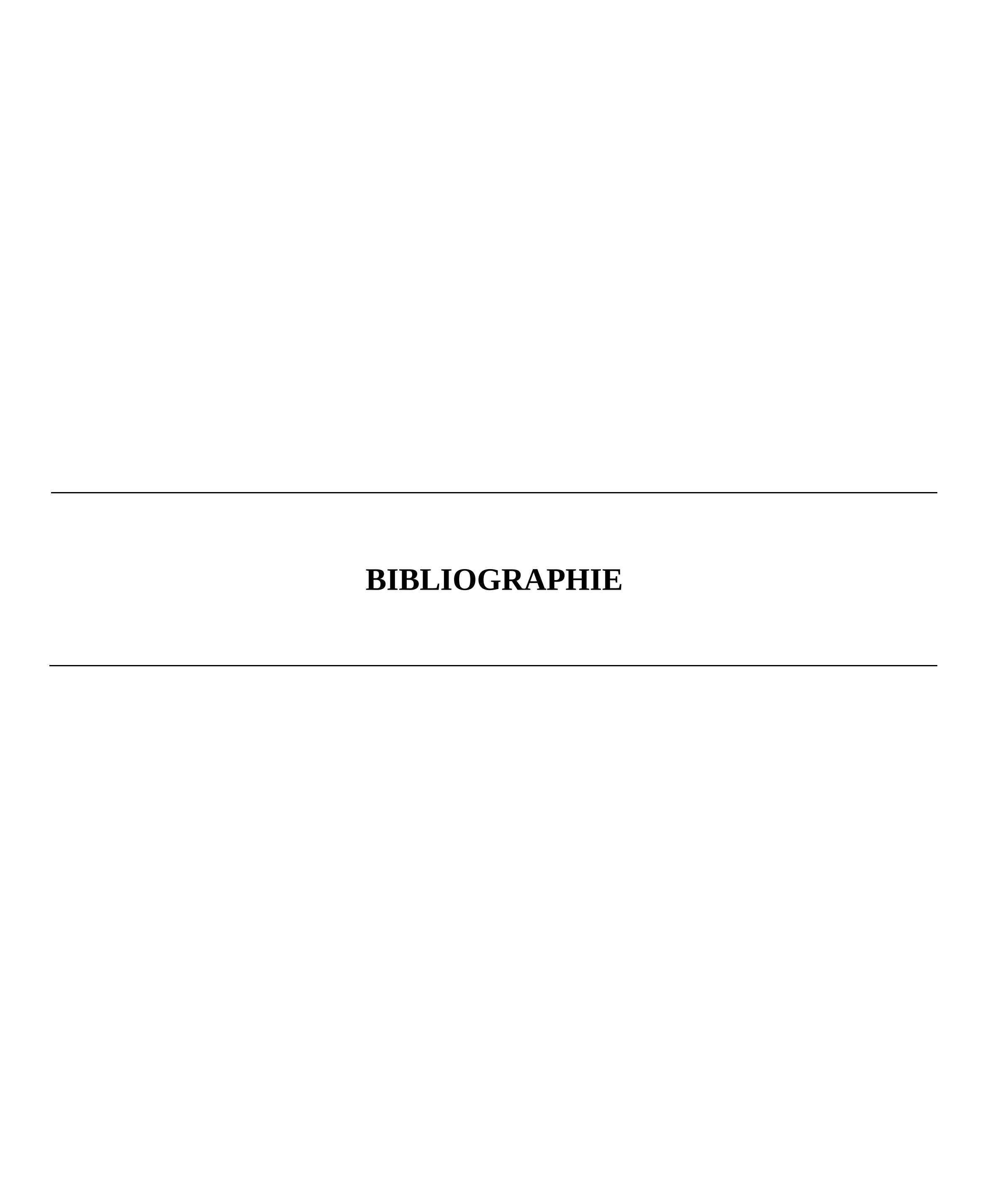
Si nous voulons résoudre le problème des unités polysémiques, nous devons prendre en considération les deux facteurs qui nous permettront la désambiguïsation sémantique : le contexte et le cotexte. Or, le contexte est un fait dynamique qui change d'une situation de communication à une autre. Il est difficile de tenir compte de tous les contextes où l'unité polysémique apparait.

Mais, le cotexte est un élément fixe dans toutes les langues naturelles. Le cotexte est relié à la syntaxe d'une langue. Nous avons vu dans notre travail les différents problèmes liés à la syntaxe que Google a rencontrés lors de la traduction des deux articles de notre corpus. Nous avons aussi expliqué comment le sens des unités polysémiques se construit par rapport à

ses relations avec les autres unités qui l'entourent. Pour cela, nous dirons que le premier problème qu'il faut résoudre en traduction automatique des unités, en l'occurrence des unités polysémiques, est celui de la syntaxe.

Ainsi, il faut mener Google, en premier lieu, à ce qu'il soit capable de bien segmenter les énoncés. Par exemple mettre des indices mathématiques pour qu'il puisse reconnaitre que tel mot commence en un point précis et s'arrête à un autre bien déterminé. Puis, il faut régler le problème de la syntaxe, c'est-à-dire essayer de bien déterminer les spécificités des syntaxes de toutes les langues dans la base de données de Google.

Nous concluons notre travail en affirmant que la polysémie n'est que l'un des problèmes de sens qui posent des difficultés à Google. Il existe d'autres phénomènes linguistiques qui posent le même problème de désambiguïsation du sens tel les homonymes. Mais, tous ces phénomènes linguistiques seraient mieux traités automatiquement si le problème de syntaxe était résolu.



BIBLIOGRAPHIE

OUVRAGES

ABEILLÉ, A. Sciences *cognitives : Sémantique pour l'analyse de la linguistique à l'informatique*. Paris, Ed Masson, 1997.

ADAM, J.M. *Linguistique et discours littéraire : théorie et pratique des textes*. Larousse université, Canada, 1976.

BARTHES, R. *Éléments de sémiologie, Communications, 4*. Paris, Seuil, 1964.

BAYLON, CHR & FABRE, P. *La sémantique*. Paris, Nathan, 1978.

BELLOS, David. *Le poisson et le bananier : Une histoire fabuleuse de la traduction.* Flammarion, Paris, 2012.

BENVENISTE, E. *Problèmes de linguistique générale I*. Paris, Gallimard, 1966.

BENVENISTE, E. *Problèmes de linguistique générale II*. Ed Gallimard, Paris, 1974.

BERSINI, H. *De l'intelligence humaine à l'intelligence artificielle*. Paris, Ellipses, 2006.

BIANCHI, CL. *La flexibilité sémantique : une approche critique, Langue française*. Paris, Larousse, 2001.

BOUILLON, P. *La traductique*. Montréal, Les presses universitaires de Montréal AUPELF URUF, 1993.

BRÉAL, M. *Essai de sémantique : sciences des significations*. Paris, Hachette, 1897.

De SAUSSURE, F. *Cours de linguistique générale*, Lausanne, Payot, 1995.

DOUAY, G et Roulland, D. *Les mots de Gustave Guillaume : vocabulaire technique de la psychomécanique du langage*. Rennes, Presses Universitaires Rennes II, 1990.

DREYFUS, H. *Intelligence artificielle, mythes et limites*. Paris, Flammarion, 1984.

ECO, UMBERTO. *Expériences de traduction*. Paris, Grasset, 2007.

ELUARD, R. *Pour aborder la linguistique*. Paris, ESF, 1993.

FREGE, G. *Sens et Référence*. Revue Zeitschrift für Philosophie und philosophische Kritik, 1892.

GEVAUDAN, P. *La polysémie verticale : Hypothèses, analyses et interprétations*. PhiN. Philologie im Netz2/a997.

GOSSELIN, L. *Le traitement de la polysémie contextuelle dans le calcul sémantique*. Intellectica, 1996.

GUILLAUME, G. *Langage et science du langage.* Paris, Librairie Nizet, 1994.

GUILLAUME, G. *Principes de linguistique théorique*. Paris, Klincksieck, 1983.

GUILLAUME, P. *La psychologie de la forme*, Paris, Flammarion,1979.

HARRIS, R. *Language, Saussure and Wittgenstein. How to play games with words*. London, Routledge, 1988.

Hellal,Y. *La théorie de la traduction: approche thématique et pluridisciplinaire*. Alger, office des Publications Universitaires, 2001.

KLEIBER, G. *Problèmes de sémantique : la polysémie en questions*. Paris, Presses Universitaires du Septen-Trion, 1999.

KLEIBER, G. *Polysémie et référence : la polysémie, un phénomène pragmatique ?*. Cahier de lexicologie n 44, 1984.

KRISTEVA, J. *Comment parler à la littérature, Tel Quel*. , paris, Seuil1968.

MEL'CUK. *Vers une linguistique Sens-Texte*. Leçon inaugurale au collège de France, Chaire internationale, Les presses de l'Université de Montréal, 1997.

MOREAU, M.L. *Sociolinguistique : concepts de base*. Paris, Mardaga, 1997.

PICOCHE, J. *Précis de lexicologie française*. Paris, Nathan, 1977.

PICOCHE, J. *Structures sémantiques du lexique français*. Paris,Nathan, 1986.

PICOCHE, J. *Polysémie n'est pas ambigüité, Cahiers de praxématique*. Université Paul Montpelier, Valéry, 1989.

RASTIER, F. *Représentation du contenu lexical et formalismes de l'Intelligence Artificielle*. Langages, 1987.

RIEGEL, M. et Pellat, J-C. et Rioul, R. *Grammaire méthodique du français*. Pais, PUF, 1994.

POTTIER, B. *Linguistique Générale ; théorie et description*. Paris, Klincksieck, 1974.

RÉCANATI, F. *La polysémie contre le fixisme. Langue Française*. N 113, Mars 1997. Aux sources de la polysémie nominale.

SABAH, G. *L'intelligence artificielle et le langage*. Hermès, 1989.

SANDERS PEIRCE, Charles. *Collected Papers*. 8 vol. Cambridge, Harvard U.P., 1931-1958. *Writings of Charles S. Pierce: A Chronological Edition*. Bloomington, Indiana U. P., 1982-1989.

TAMBA-MECSZ, I. *Le sens figuré*. Paris, Presses universitaires de France, 1981.

VISETTI, Y.M. *Modèles connexionnistes et représentation structurées*. Intellectica, 1990.

WAGNER, R.L et Pichon, J. *La grammaire du français classique et moderne*, Hachette, 1962.

DCITIONNAIRES

DUBOIS, J. *Dictionnaire de linguistique et des sciences du langage* Paris. Larousse. 1994.

GREIMAS, A.J et COURTÉS, J. *Sémiotique Dictionnaire raisonné de la théorie du langage*. Paris, Hachette, 1979.

Le nouveau petit Robert de la langue française. Paris, 2oo8, 2592.

Le Petit Larousse illustré. Paris, Larousse, 1991.

WEBOGRAPHIE

http://books.google.fr/books?id=1umrXw4cz78C&pg=PA49&dq=s%C3%A9miotique+subjectale+coquet+traduction&hl=fr&ei=azrlTYqhM8ju0gGP4yfBw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC8Q6AEwAA#v=onepage&q=s%C3%A9miotique%20subjectale%20coquet%20traduction&f=false

http://translate.google.com/

http://www.lequotidien-oran.com/

http://www.lequotidien-oran.com/index.php?news=5158631 15 octobre 2012.

http://www.lequotidien-oran.com/index.php?news=5174162 15 octobre 2012.